Quantum Analog Computing

Michail Zak

Ultracomputing Group, Mail Stop 126-347
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

Abstract

Quantum analog computing is based upon similarity between mathematical formalism of quantum mechanics and phenomena to be computed. It exploits a dynamical convergence of several competing phenomena to an attractor which can represent an extrenum of a function, an image, a solution to a system of ODE, or a stochastic process. In this paper, a quantum version of recurrent neural nets (QRN) as an analog computing device is discussed. This concept is introduced by incorporating classical feedback loops into conventional quantum networks. It is shown that the dynamical evolution of such networks, which interleave quantum evolution with measurement and reset operations, exhibit novel dynamical properties. Moreover, decoherence in quantum recurrent networks is less problematic than in conventional quantum network architectures due to the modest phase coherence times needed for network operation. Application of QRN to simulation of chaos, turbulence, NP-problems, as well as data compression demonstrate computational speedup and exponential increase of information capacity.

1. Introduction

Analog computing is based upon similarity between mathematical formalism of a physical phenomenon to be simulated and phenomena to be computed. Usually it exploits a dynamical convergence of a physical process to a certain state, or attractor, so that the measured parameters characterizing this attractor can be uniquely identified. Thus, unlike digital computers which operate via manipulations with numbers, in analog computers numbers appear as a result of measurement of physical parameters. That is why the criteria of computational complexity developed for digital algorithms, strictly speaking, are not applicable to analog algorithms. At the same time, analog algorithms have their own criteria of "complexity" such as: the time of convergence to an attractor subject to a prescribed error, the degree of stability of the attractor, the pattern of convergence (asymptotic, or oscillatory), type of the attractor (static, periodic, chaotic, or stochastic), etc.

The competition between digital and analog computers, i.e., between computations and simulations, has a long history. During the last fifty years, the theory of computation has been based, implicitly, upon classical physics as idealized in the deterministic Turing machine model. However, despite the many successes of digital computers, the existence of so called hard problems has revealed limitations on their capabilities, since the computational time for solving such problems grows exponentially with the size of the problem.

It was well understood that one possible way to fight the "curse" of the combinatorial explosion is to enrich digital computers with analog devices. In contradistinction to a digital computer, which performs operations on numbers symbolizing an underlying physical process, an analog computer processes information by exploiting physical phenomena directly. It is this problem solving via direct simulation that allows an analog approach to reduce the complexity of the computations significantly. This idea was stressed by Feynman [1] who demonstrated that the problem of exponential complexity in terms of calculated probabilities can be reduced to a problem of polynomial complexity in terms of simulated probabilities. Conceptually, a similar approach can be applied to the whole class of NP-complete problems. Indeed, the theory of computational complexity is an attribute of digital approach to computations. At the same time, in principle, one can find such a physical phenomenon whose mathematical description is equivalent to those of a particular NP-complete problem. Then, incorporating this phenomenon into an appropriate analog device, one can simulate the corresponding NP-complete problem. But is it possible, in general, to find a new mathematical formulation for any intractable problem in such a way that it becomes tractable? Some experts in computational complexity believe that, in the spirit of the Godel theorem, there always exists computational problems such that every mathematical formulation that captures the essence of the problem is intractable [2]. At this step, we cannot prove or disprove this statement.

There is another class of problems for which simulations are superior over computations. In contradistinction to NP-complete problems whose complexity is in an exponentially large number of simple computations, these problems have complex and sometimes, partially unknown analytical structure. Simulations of solutions to such problems are based upon a black-box approach when unknown components of the model are found in the course of a trial-and-error learning process. A typical representative of a corresponding analog device implementing black-box based simulations is a neurocomputer where unknown (learnable) parameters are incorporated in the form of synaptic interconnections between dynamical units called "neurons". However, usually analog computers are associated with certain limitations such as the lack of universality, slow

performance, and low accuracy, and this is the price to be paid for certain advantages of simulations. A partial success in development of a universal analog device is associated with neurocomputers which are based upon massively parallel adaptive dynamical systems modeled on the general features of biological neural networks that are intended to interact with the object of the real world in the same way the biological systems do. However, the capacity of the neurocomputers is roughly proportional to the size of the apparatus, and that limits actual power significantly.

A second way to fight a curse of dimension is to utilize a non-deterministic approach to computations. This approach is associated with the Monte Carlo method introduced by N.C. Metropolis and S.M. Ulam in 1940. The idea of this method is based upon the relationships between the probabilistic characteristics of certain stochastic processes and solutions to some deterministic problems such as values of integrals, solutions to differential equations, etc. The strength of the method is that its error does not depend on the number of variables in the problem, and therefore, if applicable, it breaks the curse of dimension. The effectiveness of the Monte-Carlo approach is inversely proportional to the smoothness parameter that characterizes the degree of correlation within the input data. However, the Monte-Carlo method is not the only way to apply nondeterminism for computations. There is a class of so-called randomized algorithms that are effective for combinatorial problems. In general, a randomized strategy for this kind of problem is useful when there are many ways in which an algorithm can proceed, but it is difficult to determine a way that is guaranteed to be good. In particular, if the benefits of good choices outweighs the costs of bad choices, a random selection of good and bad choices can yield a good algorithm.

In general, the theory of computational complexity proves that polynomial time nondeterministic algorithms are more powerful than polynomial time deterministic ones. However, the main limitation of the whole non-deterministic approach is in the generation of random numbers: the generators are slow and not always reliable (i.e., the sequence of numbers that they produce may harbor hidden correlations that no truly random sequence would possess). That is why the concept of a quantum computer became so attractive: its analog nature is based upon physical simulations of quantum probabilities, and at the same time, it is universal (at least for modeling physical world).

Although the development of the quantum-mechanical device is still in progress, a new quantum theory of computations has been founded [3]. This theory suggests that there is a second fundamental advantage of the hypothetical quantum computer which is based upon the wave properties of quantum probabilities: a single quantum computer can follow many distinct computational paths all at the same time and produce a final output

depending on the interference of all of them. This particular property opened up a new chain of algorithms that solve in polynomial time such hard problems as factorization and discrete log, i.e., the problems that are believed to be intractable on any classical computer.

There are remarkably few (actually four) papers in which quantum analog computing is discussed. The first one [4] introduces a hypothetical quantum device (a slot machine) for solving a traveling salesman problem. As shown by the author, such a device, although intellectually appealing, requires an exponentially large number of measurements to get the right answer. In the second one [5], an attempt is made to exploit combinatorial properties of tensor product decomposability of unitary evolution of many-particle quantum systems for simulating solutions to NP-complete problems, the reinforcement and selection of the desired solution being executed by quantum resonance; although the implementability of the approach is still in question, the potential difficulties are not associated with the NP-completeness of the problem. The last publication [6] introduces a new dynamical paradigm: quantum recurrent nets (QRN) which will be discussed in this paper.

The rest of the paper is organized as follows. In the next section, a QRN as an analog quantum device is introduced. Then, in Section 3, the dynamical complexity of QRN is demonstrated and discussed. Sections 4 and 5 discuss the computational power of QRN with applications to combinatorial optimization, and to simulations of chaos and turbulence. Section 6 addresses information processing by QRN; special attention is paid to data compression and recognition of collection of patterns. In Section 7, QRN as a generator of stochastic processes is discussed. The results are summarized in Section 8.

2. Quantum Recurrent Nets

Quantum recurrent nets (QRN) were introduced and discussed in [6]. They are represented by quantum version of recurrent neural networks whose outputs are coupled to their inputs via measurements and reset operations.

2.1 The Simplest QRN

The simplest QRN is described by the following set of difference equations with constant time delay τ

$$a_i(t+\tau) = \sigma_1\left\{\sum U_{ij}(t)a_j(t)\right\}, \; i.e., \left\{a_o a_{1\ldots} a_n\right\} \rightarrow \left\{0,0\ldots\underset{\uparrow_i}{1}\ldots 00\right\} \qquad (1)$$

$$i = 1,2\ldots n$$

where a_j is the input to the network at time t, U_{ij} is a unitary operator defined by the corresponding Hamiltonian of the quantum system, and σ_1 is a measurement operator (in the computational basis) that has the effect of projecting the evolved state into one of the eigenvectors of σ_1. The curly brackets are intended to emphasize that σ_1 is to be taken as a measurement operation with the effect similar to those of a sigmoid function in classical neural networks (Fig. 1). Obviously, the outputs $a_i(t+\tau)$ are random because of the probabilistic nature of quantum measurements. As shown in [6], these outputs form a Markovian stochastic process with the probabilities evolving according to the following chain:

$$\pi_i(t+\tau) = \sum_{j=1}^{n} \pi_j(t) p_{ij}, \quad \sum_{i=1}^{n} \pi_i = 1, \quad \pi_i \geq 0, \quad i = 1,2\ldots n \tag{2}$$

where $\pi = \pi(x,t);\ x = 1,2,\ldots n;\ \ \pi_i = \pi(x_i,t)$ is the n-dimensional probability vector, and

$$p_i^j = p_{ij} = \left|U_{ji}\right|^2, \qquad \sum_{j=1}^{n} p_{ij} = 1, \qquad p_{ij} \geq 0, \quad i,j = 1,2,\ldots n \tag{3}$$

is the nxn stochastic matrix which is uniquely defined by the unitary matrix U. Each element of these matrix represents the probability that the i^{th} eigenvector as an input produces j^{th} eigenvector as an output:

$$\left\{00 \quad \underset{\uparrow}{0}10 \quad 0\right\} \rightarrow \left\{00 \quad \underset{\uparrow}{0}10 \quad 0\right\} \tag{4}$$

In a special case when [7]

$$p_{ij} > 0; \quad i,j = 1,2,\ldots n \tag{5}$$

the Markov process is ergodic, i.e., the solution to Eq. (2) approaches an attractor

$$\pi \rightarrow \pi^{\infty} \ \ at\ t \rightarrow \infty \tag{6}$$

which is unique and it does not depend upon the initial value π_0 at t=0. Only this case will be considered in this paper. Thus, Eq. (1) describes the evolution of the vector

$$\{a_1 \ldots a_n\} = <\varphi|, \quad \sum_{j=1}^{n} a_j^2 = 1 \tag{7}$$

representing a quantum state in a Hilbert space, and all the components $\left(a_i U_{ij}\right)$ are to be actually implemented (Fig. 2a). This evolution is irreversible, nonlinear and nondeterministic because it includes measurement operations.

On the other hand, the vector

$$\left(\pi_1, \pi_2 \ldots \pi_n\right) = \pi, \quad \sum_{j=1}^{n} \pi_j = 1, \quad \pi_i > 0, \tag{8}$$

as well as the stochastic matrix p_{ij} exist only in an abstract Euclidean space: they never appear explicitly in physical space. The evolution (2) is also irreversible, but unlike (1), it is linear and deterministic. (Fig. 2b).

The only way to reconstruct the probability vector $\pi(t)$ is to utilize the measurement results for the vector a(t). In general case, many different realizations of Eq. (1) are required for that purpose. However, if the condition (5) holds, the ergodic attractor $\pi = \pi^{\infty}$ can be found from the only one realization due to the ergodicity of the stochastic process. The ergodic attractor π^{∞} can be found analytically from the steady-state equations:

$$\pi_i^{\infty} = \sum_{j=1}^{n} p_{ij} \pi_j^{\infty}, \quad \sum_{i=1}^{n} \pi_i^{\infty} = 1, \quad \sum_{j=1}^{n} p_{ij} = 1, \qquad \pi_i > 0, \; p_{ij} > 0 \tag{9}$$

This system of n+1 equations with respect to n unknowns $\pi_i^{\infty} (i = 1,2,\ldots n)$ has a unique solution.

As an example, consider a two-state case (n=2):

$$p_{11}\pi_1^{\infty} + p_{21}\pi_2^{\infty} = \pi_1^{\infty}, \quad p_{12}\pi_1^{\infty} + p_{22}\pi_2^{\infty} = \pi_2^{\infty} \tag{10}$$

Utilizing the constraints in Eqs. (9) one obtains:

$$p_{11}\pi_1^\infty + (1-p_{22})(1-\pi_1^\infty) = \pi_j^\infty, \quad \pi_1^\infty(p_{11}+p_{22}-2) = p_{22}-1 \tag{11}$$

whence

$$\pi_1^\infty = \frac{1-p_{22}}{2-(p_{11}+p_{22})}, \quad \pi_2^\infty = \frac{1-p_{11}}{2-(p_{11}+p_{22})} \tag{12}$$

while $$p_{11} = |u_{11}|^2, \qquad p_{22} = |u_{22}|^2 \tag{13}$$

Hence on the first sight, there are infinite number of unitary matrices u_{ij} which provide the same ergodic attractor (12). However, such a redundancy is illusive since the fact that the stochastic matrix p_{ij} has been derived from the unitary matrix u_{ij} impose a very severe restriction upon p_{ij}: not only the sum of each row, but also the sum of each column is equal to one, i.e., now in addition to Eqs. (10):

$$\sum_{i=1}^{n} p_{ij} = 1 \tag{14}$$

This makes the matrix $\|p_{ij}\|$ doubly stochastic which always leads to an ergodic attractor with uniform distribution of probabilities. Obviously such a property significantly reduces the usefulness of the QRN. However, as will be shown below, by slight change of the QRN architecture, the restriction (13) can be removed.

2.2 Multivariate QNR

In the previous sub-section we have analyzed the simplest quantum neural net whose probabilistic performance was represented by a single-variable stochastic process equivalent to generalized random walk. In this section we will turn to multi-variable stochastic process and start with the two-measurement architecture. Instead of (14) now we have the following mapping:

$$\frac{1}{\sqrt{2}}\left\{00\ldots\underset{\uparrow}{1}0\ldots\underset{\uparrow}{1}0\ldots0\right\}\rightarrow\frac{1}{\sqrt{2}}\left\{00\ldots\underset{\uparrow}{1}0\ldots\underset{\uparrow}{1}0\ldots0\right\} \tag{15}$$

$i_1 \quad i_2 \qquad\qquad j_1 \quad j_2$

i.e., $$I_1 + I_2 \rightarrow J_1 + J_2 \tag{16}$$

where $I_1, I_2, J_1,$ and J_2 are the eigenstates with the unit 1 is at the $i_1^{th}, i_2^{th}, j_1^{th}$ and j_1^{th} places, respectively. then the transitional probability of the mappings:

$$p_{i_1 i_2}^{i_1}\left(I_1 + I_2 \rightarrow J_1\right) = \frac{1}{2}\left|\overset{*}{J}_1 U\left(I_1 + I_2\right)\right|^2 = \frac{1}{2}\left|U_{j_1 i_1} + U_{j_1 i_2}\right|^2 \tag{17}$$

$$p_{i_1 i_2}^{i_2}\left(I_1 + I_2 \rightarrow J_2\right) = \frac{1}{2}\left|\overset{*}{J}_2 U\left(I_1 + I_2\right)\right| = \frac{1}{2}\left|U_{j_2 i_1} + U_{j_2 i_2}\right|^2 \tag{18}$$

Since these mapping result from two independent measurements, the joint transitional probability for the mapping (15) is

$$= p_{i_1 i_2}^{j_1 j_2}\left(I_1 + I_2 \rightarrow J_1 + J_2\right) = p_{i_1 i_2}^{j_1}\ p_{i_1 i_2}^{j_2} =$$

$$= \frac{1}{4}\left|U_{j_1 i_1} + U_{j_1 i_2}\right|^2 \left|U_{j_2 i_1} + U_{j_2 i_2}\right|^2 \tag{19}$$

One can verify that

$$\sum_{j=1}^{n} p_{i_1 i_2}^{j} = 1, \qquad \sum_{j_1 j_2 = j}^{n} p_{i_1 i_2}^{j_1 j_2} = 1 \tag{20}$$

It should be emphasized that the input patterns I_1 *and* I_2 interfere, i.e., their probabilities are added according to the quantum laws since they are subjected to a unitary transformation in the quantum device. On the contrary, the output patterns J_1 and J_2 do not interfere because they are obtained as a result of two independent measurements.

As mentioned above, Eq. (19) expresses the joint transition probabilities for two stochastic processes

$$I_1 \to J_1 \quad \text{and} \quad I_2 \to J_2 \tag{21}$$

which are coupled via the quantum interference (17) and (18)

$$I_1 + I_2 \to J_1 + J_2 \tag{22}$$

At the same time, each of the stochastic processes (21) considered separately has the transition probabilities following from Eq. (3):

$$I_1 + I_2 \to J_1 + J_2 \tag{23}$$

and by comparing Eqs. (19) and Eq. (23), one can see the effect of quantum interference for input patterns.

It is interesting to notice that although the probabilities in Eqs. (19) and (23) have a tensor structure, strictly speaking they are not tensors. Indeed, if one refers the Hamiltonian H, and therefore the unitary matrix U to a different coordinate system, the transformations of the probabilities (15) and (23) will be different from those required for tensors. Nevertheless, one can still formally apply the chain rule for evolution of transitional probabilities, for instance:

$$p_{i_1 i_2}^{q_1 q_2}\left(I_1 + I_2 \to J_1 + J_2 \to Q_1 + Q_2\right) = p_{i_1 i_2}^{q_1 q_2}\ p_{j_1 j_2}^{q_1 q_2} \quad etc. \tag{24}$$

Eqs. (19) and (24) are easily generalized to the case of ℓ measurements $(\ell \le n)$:

$$p_{i_1 \ldots i_\ell}^{j_1 \ldots j_\ell} = \frac{1}{\ell^\ell} \prod_{\alpha=1}^{\ell} \left| \sum_{\beta=1}^{\ell} U_{j_\alpha i_\beta} \right|^2, \qquad p_{i_1 \ldots i_\ell}^{q_1 \ldots q_\ell} = p_{i_1 \ldots i_\ell}^{j_1 \ldots j_\ell}\ p_{j_1 \ldots j_\ell}^{q_1 \ldots q_\ell} \tag{25}$$

Now the evolution in physical space, instead of Eq. (1), is described by the following:

$$\alpha_i(t+\tau) = \sigma_\ell \left\{ \sum U_{ij} \alpha_i(t) \right\},\ i = 1,2,\ldots n \tag{26}$$

where σ_ℓ is the ℓ-measurements operator.

Obviously, the evolution of the state vector α_1 is more "random" than those of Eq. (1) since the corresponding probability distribution depends upon ℓ variables.

2.3 QRN with input interference

In order to remove the restriction (14), let us turn to the architecture shown in Fig. 4 and assume that the result of the measurement, i.e., a unit vector $a_m(t) = \left\{00\ldots0\underset{\uparrow}{1}0\ldots0\right\}$ is combined with an arbitrary complex (interference) vector.

If the interference state vector is

$$|\psi'\rangle = \begin{pmatrix} a'_0 \\ a'_1 \\ \vdots \\ a'_N \end{pmatrix} \tag{27}$$

and σ is a measurement operator in the computational basis, then $|\psi(t+\Delta t)\rangle$, the recurrent state re-entering the circuit, must take one of the forms:

$$|\phi_0\rangle = \frac{1}{\sqrt{R_0}} \begin{pmatrix} 1+a'_0 \\ a'_1 \\ \vdots \\ a'_{N-1} \end{pmatrix} = \frac{1}{\sqrt{R_0}} \begin{pmatrix} a_0^{(0)} \\ a_1^{(0)} \\ \vdots \\ a_{N-1}^{(0)} \end{pmatrix}$$

$$|\phi_1\rangle = \frac{1}{\sqrt{R_1}} \begin{pmatrix} a'_0 \\ 1+a'_1 \\ \vdots \\ a'_{N-1} \end{pmatrix} = \frac{1}{\sqrt{R_1}} \begin{pmatrix} a_0^{(1)} \\ a_1^{(1)} \\ \vdots \\ a_{N-1}^{(1)} \end{pmatrix} \tag{28}$$

$$|\phi_{N-1}\rangle = \frac{1}{\sqrt{R_{N-1}}} \begin{pmatrix} a'_0 \\ a'_1 \\ \vdots \\ 1+a'_{N-1} \end{pmatrix} = \frac{1}{\sqrt{R_{N-1}}} \begin{pmatrix} a_0^{(N-1)} \\ a_1^{(N-1)} \\ \vdots \\ a_{N-1}^{(N-1)} \end{pmatrix}$$

with re-normalization factors:

$$R_0 = |1+a'_0|^2 + |a'_1|^2 + \ldots$$

$$R_1 = |a_0'|^2 + |1 + a_1'|^2 + \ldots \tag{29}$$
$$\vdots$$
$$R_{N-1} = |a_0'|^2 + |a_1'|^2 \ldots + |1 + a_{N-1}'|^2$$

The transition probability matrix, p_i^j for this process is given by examining how each of the recurrent states, $|\phi_0\rangle \cdots |\phi_{N-1}\rangle$ evolve under the action of U:

$$p_i^j = \begin{pmatrix} \left|\frac{b_0^{(0)}}{\sqrt{R_o}}\right|^2 & \left|\frac{b_1^{(0)}}{\sqrt{R_0}}\right|^2 & \cdots \\ \left|\frac{b_0^{(1)}}{\sqrt{R_1}}\right|^2 & \left|\frac{b_1^{(1)}}{\sqrt{R_1}}\right|^2 & \cdots \\ \vdots & \vdots & \ddots \\ \left|\frac{b_0^{(N-1)}}{\sqrt{R_{N-1}}}\right|^2 & \cdots & \left|\frac{b_{N-1}^{(N-1)}}{\sqrt{R_{N-1}}}\right|^2 \end{pmatrix} \tag{30}$$

where

$$b_j^{(i)} = \sum_{l=0}^{N-1} U_{jl} a_l^{(i)} = U_{ji} + \sum_{l=0}^{N-1} U_{jl} a_l(0) \tag{31}$$

Thus, now the structure of the transition probability matrix p_i^j can be controlled by the interference vector (27), and in particular, the restriction (14) can be removed.

Next we generalize the concept of a quantum recurrent network to the case in which there are k networks working in parallel (see Fig. 5). During the quantum evolution and measurement phases each network acts independently of the rest. However, during the reset operation the results of all the measurements are combined with the initial state to yield k identical input states.

Note that the reset operation does not require an arbitrary quantum state to be cloned (which is impossible). Instead it only requires that k classical states, the outcomes of the k independent measurements, be copied. As this information is purely classical it can be copied freely. Moreover, the initial state $|\psi(0)\rangle$ is known and can be generated afresh as needed by each of the k networks. As a result, the feedback process we propose is guaranteed to be physically realizable.

The purpose of moving to the k-parallel quantum recurrent network is to permit us to generate *multi-dimensional* stochastic attractors of the most general form. The reset

operation which includes the interference input gives us the flexibility to introduce correlations between the attractors in each dimension.

For the k-parallel QRN, the elements of the transition probability matrix now define the probability of making transitions between sets of measurement outcomes. The state entering each of the k networks at each iteration will have a form such as

$$\left|\phi_{i_1 i_2 \cdots j_k}\right\rangle = \frac{1}{\sqrt{R_{(i_1 i_2 \ldots i_k)}}} \begin{pmatrix} a_0^{(i_1 i_2 \cdots i_k)} \\ a_1^{(i_1 i_2 \cdots i_k)} \\ \vdots \\ a_{N-1}^{(i_1 i_2 \cdots i_k)} \end{pmatrix} \tag{32}$$

where the sequence $i_1 i_2 \ldots i_k$ specifies the last ordered set of measurement outcomes obtained from the k networks and $R_{i_1 i_2 \ldots i_k}$ is the renormalization constant given by:

$$R_{(i_1 i_2 \ldots i_k)} = \left|a_0^{(i_1 i_2 \ldots i_k)}\right|^2 + \left|a_1^{(i_1 i_2 \ldots i_k)}\right|^2 + \cdots \left|a_{N-1}^{i_1 i_2 \ldots i_k}\right|^2$$

The mathematical form of the amplitude $a_i^{(i_1 i_2 \ldots i_k)}$ depends upon how many of the components in the k-parallel network produced the same measurement outcome at the last iteration through the QRN i.e., how many of the i_i in the sequence $i_1 i_2 \ldots i_k$ were the same. If the outcome i_i is obtained n_{i_i} times we have a $a_i^{(i_1 i_2 \ldots i_k)} = n_{i_i} + a_i(0)$.

As there are k networks and each network can produce one of N outcomes (independently), the k-parallel transition matrix defines a mapping from N^k distinct sets of input states to N^k sets of output states. If we denote the probability of transitioning from the set of inputs $i_1 i_2 \ldots i_k$ to the set of outputs $j_1 j_2 \ldots j_k$ by $p_{j_1 j_2 \ldots j_k}^{(i_1 \ldots i_k)}$ we have:

$$p_{j_1 \ldots j_k}^{(i_1 \ldots i_k)} = \left|\frac{b_{j_1(U_1)}^{(i_1 i_2 \ldots i_k)}}{\sqrt{R_{(i_1 i_2 \ldots j_k)}}}\right|^2 \left|\frac{b_{j_2(U_2)}^{(i_1 i_2 \ldots i_k)}}{\sqrt{R_{(i_1 i_2 \ldots i_k)}}}\right|^2 \left|\frac{b_{j_k(U_k)}^{(i_1 i_2 \ldots i_k)}}{\sqrt{R_{(i_1 i_2 \ldots i_k)}}}\right|^2 \tag{33}$$

where

$$b_{j(U)}^{(i_1 i_2 \ldots i_k)} = \sum_{s=1}^{k} U_{j i_s} + \sum_{t=0}^{N-1} U_{j t} a_t(0)$$

thus the k-parallel transition probability matrix has a tensor structure of the form $T_k = \left\{ p_{j_1 j_2 \ldots j_k}^{(i_1 i_2 \ldots i_k)} \right\}_{N^k \times N^k}$ where the sequences $i_1 i_2 \ldots i_k$ and $j_1 j_2 \ldots j_k$ are defined with respect to some consistent ordering.

For the k-parallel architecture, there are kN^2 free parameters. Thus we ought to be able to generate k-dimensional stochastic attractors having up to kN^2 degrees of freedom. If all the unitary matrices in (33) are the same:

$$U_1 = U_2 = \ldots = U_k$$

one arrives at the ℓ-measurement architecture where Eq. (33) generalizes Eq. (25) when the input interference is applied.

In order to clarify the more complex ℓ-measurement architecture of QRN, for instance, such as those given by Eqs. (26), turn to Eq. (19).

By simple manipulation of indices one obtains:

$$\pi_{i_1 i_2}(t+\tau) = \sum_{j_1, j_2 = 1}^{n} \pi_{j_1 j_2}(t) p_{i_1 i_2}^{j_1 j_2}, \qquad \sum_{i_1 i_2 = 1}^{n} \pi_{i_1 i_2} = 1, \left(\pi_{i_1 i_2} \geq 0\right), \; i_1, i_2 = 1, 2, \cdots n \qquad (34)$$

Here $\pi = \pi(x_1, x_2); \; x_1, x_2 = 1, \ldots n$ is the n^2-dimensional joint probability vector, and $\left\| p_{i_1 i_2}^{j_1 j_2} \right\|$ is the $n^2 \times n^2$ stochastic matrix which is uniquely defined by the unitary matrix U and the interference vector (27), (see Eqs. (33) for $k = 2, U_1 = U_2 = \ldots = U$). Each element of this matrix represents the joint probability that the combined (normalized) input of i_1^{th} and i_2^{th} eigenvectors produces j_1^{th} and j_2^{th} eigenvectors as outputs of the first and second measurements, respectively.

Obviously

$$\sum_{j_1 j_2=1}^{n} p_{i_1 i_2}^{j_1 j_2} = 1, \qquad p_{i_1 i_2}^{j_1 j_2} \geq 0 \tag{35}$$

In a special case similar to (5) when

$$p_{i_1 i_2}^{j_1 j_2} > 0, \qquad i, j = 1,2 \ldots n \tag{36}$$

the stochastic process approaches an ergodic attractor represented by an ergodic Markov process. Since the system (34) is linear, this attractor is unique; it can be found from Eq. (34) by solving the system of n^2 linear equations with respect to n^2 components of the joint probability vector $\pi_{i_1 i_2}$:

$$\pi_{i_1 i_2}^{\infty} = \sum_{j_1 j_2=1}^{n} \pi_{j_1 j_2}^{\infty} \; p_{i_1 i_2}^{j_1 j_2} \tag{37}$$

subject to the following constraint:

$$\sum_{i_1 i_2=1}^{n} \pi_{i_1 i_2}^{\infty} = 1, \qquad \left(\pi_{i_1 i_2}^{\infty} \geq 0\right)$$

This $\left(n^2+1\right)^{th}$ equation does not over determine the system (37) because of the condition (36): as in the one-measurement case (9), the system has a unique solution. Due to the interference input (27), the probability distribution at the attractor is not necessarily uniform: its shape is controlled by this input.

It is worth emphasizing that, as follows from Eqs. (3) and (19), in general

$$p_{i_1 i_2}^{j_1 j_2} \neq p_{i_1}^{j} \otimes p_{j_2}^{j_2}$$

i.e., the two-dimensional stochastic attractor (34) is not simply the product of two underlying one-dimensional attractors.

Eqs. (34)-(37) can be generalized to ℓ-measurement architecture:

$$\pi_{i_1\ldots i_\ell}(t+\tau)=\sum_{j_1\ldots j_\ell=1}^{n}\pi_{j_1\ldots j_\ell}(t)p_{i_1\ldots i_\ell}^{j_1\ldots j_\ell},\quad \sum_{j_1\ldots j_\ell=1}^{n}p_{i_1\ldots i_\ell}^{j_1\ldots j_\ell}=1, \tag{38}$$

$$\sum_{i_1\ldots i_\ell=1}^{n}\pi_{i_1\ldots i_\ell}=1,\quad \pi_{i_1\ldots i_\ell}>0 \tag{39}$$

$$\pi_{i_1\ldots i_\ell}^{\infty}=\sum_{j_1\ldots j_\ell}^{n}\pi_{j_1\ldots j_\ell}^{\infty}p_{i_1\ldots i_\ell}^{j_1\ldots j_\ell} \tag{40}$$

$$p_{i_1\ldots i_\ell}^{j_1\ldots j_\ell}>0,\quad i,j=1,2,\ldots n \tag{41}$$

Thus, the evolution of the state vector (26) in physical space corresponds to the evolution of the probability vector $\pi_{i_1\ldots i_\ell}$ (t) in an abstract Eucledian space, while the ℓ-variate stochastic attractor of $a_i(t)$ is described by the ergodic process with the probability distribution $\pi_{i_1\ldots i_\ell}^{\infty}$ as $\pi_{i_1\ldots i_\ell}(t\to\infty)$.

It should be noticed that given the unitary matrix U, and the interference vector (27), the attractor $\pi_{i_1\ldots i_\ell}^{\infty}$ is unique; it does not depend upon initial input $a_i(0)$. As mentioned above, if the interference vector (27) is zero, then this unique attractor has a uniform probability distribution.

2.3 Non-Markovian Processes

The quantum neural nets (1) or (26), with a slight modification, can generate non-Markovian processes which are "more deterministic" because of higher correlations between values of the vector a_i at different times, i.e., between $a_i(t), a_i(t-\tau), a_i(t-2\tau)$, etc.

Indeed, let us assume that each new measurement is combined with the ℓ previous measurements (instead of ℓ repeated measurements). Then Eq. (25) will express the joint distribution of $a_i(t), a_i(t-\tau)$,.....etc.

The evolution of the probabilities is described by equation similar to (34):

$$\pi_{i_1\ldots i_\ell}(t+\tau)=\pi_{j_1\ldots j_\ell}\ (t)\ p_{i_1\cdots i_\ell}^{j_1\cdots j_\ell} \tag{42}$$

where $\pi_{i_1\ldots i_\ell}$ is the joint probability distribution for the vectors

$$a_i(t), a_i(t-\tau), a_i(t-2\tau)\ \ etc.$$

Thus, instead of ℓ-dimensional Markov process, now we have a one-dimensional non-Markovian process of the ℓ^{th} order.

3. Dynamical Complexity of QRN's

3.1 Shannon and algorithmic complexity.

Although the concept of complexity is well understood intuitively, its strict definition remains an enigma since there are many different aspects which can be associated with complexity (the number of interacting variables, the degree of instability, the degree of determinism, etc.). Here we will associate dynamical complexity with the degree of unpredictability of the underlying motion. Then the Shannon entropy becomes the most natural measure of dynamical complexity of QRNs:

$$H=-\sum_{i=0}^{n-1}\pi_1 \ell og_2 \pi_i$$

and (43)

$$H_{\max}\infty \ell og_2 n$$

Let us assume now that the unitary matrix in Eq. (1) is composed of a direct product of n 2×2 unitary matrices:

$$U=U_1\otimes U_2\otimes\ldots\otimes U_n, \qquad N=2^n=2^{q/4} \tag{44}$$

where the number of independent components in U_1

$$q=4n \tag{45}$$

Then the dynamical complexity of QRN becomes exponentially larger (see Eq. (44):

$$H_{\max} \propto \ell og_2 2^n \propto n$$

although the algorithmic complexity is still expressed by Eq. (3) or more precisely, by Eq. (45). Thus, QRNs based upon representation (41) generate "complexity" in an exponential rate, and therefore the underlying stochastic processes attain structure of fractals. Indeed, as shown in [8], a continuous version of a Markov process exhibits self-similar structure down to infinitesimal scales of observation. Although the Markov processes generated by QRNs are finite-dimensional, their scales approaches zero exponentially fast when the number of the variables n (see Eq. (41) grows only linearly. This means that QRNs generate "quantum fractals" which can be applied to image compression, animation, or for a finite-dimensional representation of Weierstrass-type functions which are continuous but non-differentiable. In contradistinction to classical fractals, quantum fractals are more controllable since their probabilistic structure can be prescribed.

Now suppose that we are interested in generating a stochastic process with prescribed probability distribution. Then the algorithmic complexity becomes important: it will allow us to preserve only q=4n (out of $N = 2^n$) independent characteristics of the distribution (although the stochastic process will be still N-dimensional, and its Shannon complexity will be of order of n).

The difference between the Shannon and the algorithmic complexities effects the design of the ℓ measurements architecture in the following way. As results from Eq. (15), the input-output relationships require the number of mapping (i.e., quantum circuits) which is polynomial in N, i.e., exponential in n. However, if the unitary matrix U has a direct-product representation (41) then, as follows from the identity:

$$(U_1 \otimes U_2) \bullet (a_1 \otimes a_2) = (U_1 a_1) \otimes (U_2 a_2) = Ua \qquad (46)$$

i.e.,

$$a = a_1 \otimes a_2 \qquad (47)$$

and therefore, not only the size of the unitary matrix U and the state vector a, but also the number of mapping circuits for ℓ-measurement architectures become polynomial in n as

far as their actual implementation is concerned. In addition to that, in the case (41) n out of ℓ measurements can be performed in parallel.

Eq. (44) is not the only representation of a unitary matrix which preserves its exponential size while utilizing only polynomial resources. Indeed, consider the following combination of direct and dot products:

$$U = \left(U_1^{(1)} \otimes \cdots U_n^{(1)}\right) \bullet \left(U_1^{(2)} \otimes \cdots U_n^{(2)}\right) \cdots \left(U_1^{(m)} \otimes \cdots U_n^{(m)}\right) \cdots \tag{48}$$

Here the number of independent components is:

$$q = 4nm \tag{49}$$

while the dimensionality

$$N = 2^n = 2^{q/4m} \tag{50}$$

In Eq. (50), N and q are associated with the Shannon and the algorithmic complexity, respectively.

3.2 Hidden complexity.

The complexity of QRNs is not exhausted by the phenomenon of quantum fractals: they can exhibit even more sophisticated behavior. Indeed, so far it was implied that QRN simulates a physical system. Now suppose that we are dealing with a biological or social dynamics when the underlying system is trying to hide its identity by intentionally misleading an observer. It turns out that QRNs possess a large capacity for simulating this type of behavior.

We will start with the case when the probability distribution characterizing the stochastic process is random itself. For that purpose, we will introduce an additional feedback from the output state which acts upon the basic Hamiltonian as follows: (Fig 3):

$$H(t+1) = H(t) + \left[a(t) \otimes a(t) + \frac{\pi}{4} I \right] \hbar \tag{51}$$

i.e.,

$$U(t+1) = U(t) e^{ia(t) \otimes a(t) + i\frac{\pi}{4}} \tag{52}$$

where a (t) is the measured output state.

Then for N=2

$$a=\begin{cases}1, \quad 0 & \text{with the probability } \pi^* \\ 0, \quad 1 & \textit{with the probability } 1-\pi^*\end{cases}$$

$$H(t+1)=H(t)+\begin{cases}\begin{pmatrix}\pi/4 & 0\\ 0 & 0\end{pmatrix} & \text{with the probability } \pi^* \\ \begin{pmatrix}0 & 0\\ 0 & \pi/4\end{pmatrix} & \text{with the probability } 1-\pi^*\end{cases}$$

$$U(t+1)=U(t)\begin{cases}\begin{pmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix} & \text{with the probability } \pi^* \\ \begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix} & \text{with the probability } 1-\pi^*\end{cases} \tag{53}$$

If $U(t)=\begin{pmatrix}U_{11} & U_{112}\\ U_{21} & U_{22}\end{pmatrix}$, then

$$U(t+1)=\frac{1}{2}\begin{cases}\begin{pmatrix}U_{11}-U_{12}, & U_{11}+U_{12}\\ U_{21}-U_{22} & U_{21}+U_{22}\end{pmatrix} & \text{with the probability } \pi^* \\ \begin{pmatrix}U_{11}+U_{12}, & U_{11}+U_{22}\\ U_{21}+U_{22} & U_{21}-U_{22}\end{pmatrix} & \text{with the probability } 1-\pi^*\end{cases} \tag{54}$$

Here the probability π^* at each time-step is defined by a chain rule: p_i^j

$$\left[\pi^*(t+1),1-\pi^*(t+1)\right]=\left[\pi^*(t),1-\pi^*(t)\right]\,P \tag{55}$$

Thus, the process described by QRN with the Hamiltonian (51) jumps randomly from one stochastic pattern to another. Eventually it may converge to a stationary ergodic process which is combined of two different stochastic processes visited randomly with time-independent probability.

In general, QRNs are capable to generate multidimensional stochastic processes which converge (with prescribed probability distribution) to several different attractors, while each of those attractors in turn, has its own probability distribution.

Now it is clear that although the physical entropy of this kind of process can be small, their complexity is much larger than of those described by the maximum entropy since they are more unpredictable: their identity is hidden behind a "double shield" of probability.

Next we will introduce analytically more trivial, but dynamically more complex processes whose probability distribution change chaotically.

Suppose that the elements of the basic unitary matrix change in time according to the following chain rule:

$$U_{11}^{(n+1)} = 2U_{11}^{(n)}\sqrt{1-\left|U_{11}^{(n)}\right|^2} = U_{22}^{(n+1)}$$

$$U_{12}^{(n+1)} = \sqrt{1-\left|U_{11}^{(n)}\right|^2} = -U_{21}^{(n+1)} \quad .$$

Then the elements of the corresponding transition probability matrix P will evolve as logistic maps:

$$p_{11}^{(n+1)} = 4p_{11}^{(n)}\left(1-p_{11}^{(n)}\right) = p_{22}^{(n+1)}$$

$$p_{12}^{(n+1)} = 1-4p_{11}^{(n)}\left(1-p_{11}^{(n)}\right) = p_{21}^{(n+1)}$$

and

$$\pi_1^{(n+1)} = \pi_1^{(n)}p_{11}^{(n)} + \pi_2^{(n)}p_{21}^{(n)} \qquad (56)$$

Since $p_{11}^{(n)}$ change chaotically (according to a logistic map) π_1 will change chaotically too.

Thus, in this case the probability distribution for the underlying stochastic process is chaotically unstable, and therefore, its true identity is hidden from an observer even deeper than in the previous case.

It should be noticed that all these complex paradigms can be exploited for simulations and analysis of dynamical processes in physics, biology and economics.

3.3 Coupled Stochastic Processes.

Let us turn to Eqs. (33) which expresses the transition probability matrix for k-parallel QRN architecture.

In order to demonstrate the power and capacity of the ℓ-parallel QRN, suppose that ℓ=2, and

$$U_1 = \begin{pmatrix} U'_{11} & U'_{12} \\ U'_{21} & U'_{22} \end{pmatrix} \neq U_2 = \begin{pmatrix} U''_{11} & U''_{12} \\ U''_{21} & U''_{22} \end{pmatrix} \tag{57}$$

If the measured output from U_1 acts as an input to U_2, and vise-versa, then one arrives at two stochastic processes whose probabilities

$$\pi_1 = \left(\pi_1^{(1)}, \pi_1^{(2)}\right) \text{ and } \pi_2 = \left(\pi_2^{(1)}, \pi_2^{(2)}\right) \tag{58}$$

are coupled:

$$\pi_1(t+\tau) = \pi_2(t) P_{12} \tag{59}$$

$$_2(t+\tau) = \pi_1(t) P_{21} \tag{60}$$

where the transition probability matrices P_{12} and P_{21} follow from Eq. (3):

$$P_{12} = \left\| |U'_{ji}|^2 \right\|, \quad P_{21} = \left\| |U''_{ji}|^2 \right\| \tag{61}$$

In this simple case Eqs. (59) and (60) can be decoupled:

$$\pi_1(t+2\tau) = \pi_1(t) P_{21} P_{12} \tag{62}$$

$$\pi_2(t+2\tau) = \pi_2(t) P_{12} P_{21} \tag{63}$$

If the resulting stochastic matrices $P_{21}P_{12}$ and $P_{12}P_{21}$ satisfy the condition (5) (which will be the case when each matrix P_{12} and P_{21} satisfies it separately) then a unique pair of coupled ergodic attractors exists, and one can find it from the system of linear equations:

$$\pi_1^\infty = \pi_1^\infty P_{21} P_{12} \tag{64}$$

$$\pi_2^\infty = \pi_2^\infty P_{12} P_{21} \tag{65}$$

Since, for simplicity, the interference input (27) was not applied, both attractors will be identical:

$$\pi_1^\infty = \left(\frac{1}{2}, \frac{1}{2}\right) = \pi_2^\infty$$

However, with the input (27) they can be different and not uniform.

3.4 Nonlinear Stochastic Processes.

So far all the stochastic processes considered above were linear. Now let us assume that along with the Eq. (1) which is implemented by quantum device, we implement (in a classical way) the associated probability equation (2). At this point, these two equations are not coupled yet. Now turning to Eqs. (27)-(33), assume that the interference vector $|\varphi'\rangle$ is played by the probability vector π. Then Eqs. (1) and (2) take the form:

$$a_i(t+\tau) = \sigma_1\left\{\sum_j U_{ij} a_i(t)\right\} \tag{66}$$

$$\pi_1(t+\tau) = \sum p_i^j \pi_j(t) \tag{67}$$

where p_i^j is found from Eqs. (28) if in the vector (27) the components a_i' are replaced by π_i, and they are coupled. Moreover, the probability evolution (64) becomes nonlinear since the matrix p_{ij} depends upon the probability vector π. Because of that, the solution to Eqs. (67) may have many different ergodic attractors (depending upon the initial conditions) or may not have ones at all.

In a particular case, Eq. (67) can be interpreted as a discretised version of a parabolic partial differential equation in which the differential operator D is replaced by the shift operator E_h :

$$E_h = \mathrm{e}^{\mathrm{hD}} \tag{68}$$

Therefore, discretised versions of such fundamental nonlinear phenomena as shock waves, Burger's waves, solitons, etc. can occur in solutions to Eq. (67). Actually these effects are associated with concentrations of probabilities, and they can be interpreted as special emerging effects of self-organization.

4. Computing by Simulations.

4.1 Exponential speedup.

As shown in Section 2, a ℓ-measurement QRN converge to an ergodic attractor which in probability space is described by an N'-dimensional probability vector with the components $\pi^{\infty}_{i_1 \cdots i_l}$. In order to find these components by computations, one has to solve a system of N' linear algebraic equations (40) subject to the constraints (39) which requires exponential resources.

At the same time, simulations of the same problem require the number of measurements which, at most, are reciprocal of the square of the error threshold $\left(1/\varepsilon^2\right)^{[2]}$ while its dependence of the dimensionality is linear: there are ℓn measurements per unit time τ are required, where n is expressed by Eq. (49) and (50), and the size of the input is nm.

In order to illustrate this statement, consider a two-measurement architecture. As follows from Eqs. (26), each pair of consecutive measurements consists of two eigenstates which can be labeled by the numbers j_1 and j_2 as indicated in Eq. (15). When the stochastic process approaches its ergodic attractor, the plane $j_1 j_2$ will be filled up by points representing pairs of the joint values $j_1^* j_2^*$ of the outputs (see Fig. 6). Actually density of these points are the values of the discretize version of the probability density $\pi_{j_1 j_2}$. However, there is a certain price to be paid for this exponential speedup: the matrix $p^{j_1 \cdots j_l}_{i_1 \cdots i_l}$ in the simulated system (40) is supposed to satisfy N' normalization constraints, and all its elements must be non-negative (see Eqs. (36) and (41). But in addition to that, the number

of independent parameters which can be varied is equal exactly to the number of independent parameters of the unitary matrix, i.e., for the case (48) it is expressed by Eq. (49)

$$q = 4nm << S = 2^{\frac{q'}{4m}} \tag{69}$$

In terms of complexity, here q corresponds to the algorithmic complexity, and S- to the Shannon complexity.

This means that the ability of mapping a desired computational problem into the system (40) is controlled by the algorithmic complexity which remains relatively low.

Now the following question can be asked: are there any practical problems which can be mapped into the specific architecture simulated by QRN?

It turns out that there is a broad class of problems described by the Fokker-Planck equation which fits exactly the simulated system, and these problems will be discussed next.

4.2 Simulation of Fokker-Planck equation

The Fokker-Planck equation deals with those fluctuations of systems which stem from many tiny disturbances, each of which changes the system in an unpredictable way. It provides a powerful tool with which the effects of fluctuations close to transition points can be adequately treated. This equation is used in a number of different fields in natural science (solid state physics, quantum optics, chemical physics, theoretical biology and circuit theory).

The Fokker-Planck equation for ℓ variables is a linear parabolic partial differential equation:

$$\frac{\partial \pi}{\partial t} = \left[-\sum_{i=1}^{\ell} \frac{\partial}{\partial x_i} D_i^{(1)}(\{x\}) + \sum_{i,j=1}^{\ell} \frac{\partial^2}{\partial x_i \partial x_j} D_{ij}^{(2)}(\{x\}) \right] \pi \tag{70}$$

with respect to the probability distribution function

$$\pi = \pi(\{x\}, t), \quad \int_{V} \pi(\{x\}, t)\, d^{\ell} x = 1 \tag{71}$$

with drift vector $D_i^{(1)}$ and the diffusion tensor $D_{ij}^{(2)}$ depending on the ℓ variables: $x_{1,\ldots}x_{\ell} = \{x\}$.

A finite-dimensional approximation to Eq. (70) obtained by replacing all the spatial derivatives with their finite-difference representations leads exactly to Eqs. (40) (including the constraints (39) imposed upon the matrix $p_{i_1\ldots i_2}^{j_1\ldots j_\ell}$) which are simulated by QRN.

It should be understood that the only stationary version of Eq. (70) can be simulated by QRN effectively because of the ergodicity advantage.

The computational complexity of integrating Eq. (70) is on the order of $(1/\varepsilon)^{\ell}$ — that is, the reciprocal of the error threshold raised to a power equal to the number of variables [10] which is exponential in ℓ. In contradistinction to that, the resources for simulations by QRN is on the order of $(1/\varepsilon^2)$, i.e., they do not depend upon the dimensionality of the problem.

There is another advantage of QRN-simulations of the Fokker-Planck equation which was predicted by Feynman [1]: suppose that we are interested in behavior of the solution to Eq. (70) in a local region of the variables $\{x\}$; then, in case of computing, one has to find the global solution first, and only after that the local solution can be extracted, while the last procedure requires some additional integrations in order to enforce the normalization constraints. On the other hand, in case of QRN simulations, one can project all the measurements onto a desired sub-space $j_\alpha \otimes j_\beta$ of the total space $j_1 \otimes \ldots j_\ell$ and directly obtain the local solution just disregarding the rest of the space.

The last comment to be made is concerning the restricted number of independent variables in the matrix $p_{1,\ldots\ell}$ of Eqs. (40) (See Eqs. (38) and (40) which is supposed to be mapped onto the diffusion coefficients $D_i^{(1)}$ and $D_i^{(2)}$. As a matter of fact, the number of non-zero components of the matrix corresponding to the discretized version of Eq. (70) will be low because of the weak coupling between variables provided by only first and second spatial derivatives: most of the far-off-diagonal components will be zero. In addition to that, the diffusion coefficients are smooth function of $\{x\}$, and therefore, they are defined by small (or, at most, polynomial) number of coefficients. Therefore, an appropriate mapping can always be performed by the corresponding increase of the number of independent variables q in Eq. (49).

4.3 Simulating optimization problems

Since Eq. (40) represents a unique attractor of the dynamical process described by Eq. (38), this attractor delivers a unique (global) minimum to the following quadratic form

$$\mathrm{L}=\sum_{i_1\ldots i_l=1}^{N}\left(\pi_{i_1\ldots i_l}-\sum_{i_1\ldots i_l=1}^{N}\pi_{i_1\ldots i_l}p_{i_1\ldots i_l}^{j_1\ldots j_l}\right)^2+\left(\sum_{i_1\ldots i_l=1}^{N}\pi_{i_1\ldots i_l}-1\right)^2 \tag{72}$$

which plays the role of a Lyapunov function for Eqs. (38) subject to the normalization constraints.

After simple algebraic transformations, Eq. (72) reduces to

$$L=\sum_{i_1\ldots i_l=1}^{N}\pi_{i_1\ldots i_l}^2\left[N^l+1-\sum_{j_1\ldots j_l=1}^{N}\left(p_{i_1\ldots i_l}^{j_1\ldots j_l}\right)^2\right]+1+$$

$$+\sum_{\substack{i_1\ldots i_l=1\\ j_1\ldots j_l=1}}^{N}\pi_{i_1\ldots i_l}\pi_{j_1\ldots j_l}\left(2N^l+1-\sum_{\substack{i_1\ldots i_l=1\\ j_1\ldots j_l=1}}^{N}p_{i_1\ldots i_l}^{k_1\ldots k_l}p_{k_1\ldots k_l}^{j_1\ldots j_l}\right)\cdots \tag{73}$$

Now it is clear that L is a positive quadratic form since

$$p_{i_1\ldots i_l}^{j_1\ldots j_l}<1 \tag{74}$$

Here N^l is the dimensionality of the probability vector.

Obviously all the advantages of quantum simulations over the computations of the minimum of L in (73) are the same as for the solution of the system (38) discussed above.

4.4 Simulating NP Problems

Let us turn to Eq. (70) and pose the following problem: find the probability $\pi\{x\}$ at a prescribed point

$$\left\{\overset{*}{x}\right\} = \overset{*}{x}_1, \ldots \overset{*}{x}_\ell \tag{75}$$

Another problem of a similar complexity is the following: find the expected value of $\{x\}$, i.e., $E\left(\left\{\overset{*}{x}\right\}\right)$ which is equivalent to finding expected values for all the coordinates, i.e.,

$$E\left(\overset{*}{x}_1\right), E\left(\overset{*}{x}_2\right), \cdots E\left(\overset{*}{x}_\ell\right) \tag{76}$$

Both of these problems are significantly different from those described above, and the difference is the following: When one is asked to find the global solution to Eqs (70) or (40), he has to deal with an exponential number of components $\pi(\{x\})$ or $\pi_{i_1 \ldots i_\ell}$, so the answer itself is exponentially "long" no matter how the solution was obtained. On the other hand, the answer to the problems (75) or (76) is very simple: it consists only of one or ℓ numbers, respectively. In addition to that, the size of the input to QRN is polynomial due to the quantum advantage: the direct product decomposability of the unitary matrix (see, for instance, Eq. (48) which contracts the 2^n-vector into mn-vectors).

The last type of problems has all the attributes of so called NP (nondeterministic polynomial) class of complexity. Now the advantage of QRN simulations over computations becomes more pronounced: QRN simulations provide exponential speedup for NP problems associated with the Fokker-Planck equation for which computations require exponential resources.

4.5 Simulating Expectations.

So far our analysis was restricted by ergodic processes enforced by the condition (5), (36) and (41). The reason for that is obvious: due to ergodicity all the information about the probability distribution can be obtained from a single run of QRN. However, there are cases when the ergodicity conditions do not hold. Indeed, consider, for instance, the simplest Fokker-Planck equation

$$\frac{\partial \pi}{\partial t} = \frac{\partial^2 \pi}{\partial x^2} \tag{77}$$

subject to reflecting boundary conditions.

A discretized version of this equation

$$\pi_{i,j+1} = \frac{\pi_{i+1,j} + \pi_{i-1,j}}{2} \tag{78}$$

is simulated by a symmetric random walk with the stochastic matrix

$$p_{ij} = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 & 0 & \cdots \\ 1/2 & 0 & 1/2 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \tag{79}$$

which does not satisfy the condition (5).

The steady state solution to Eq. (77) is trivial:

$$\pi^{\infty} = \frac{1}{L} = \text{Const} \tag{80}$$

where L is the distance between the reflecting boundaries.

The same solution can be obtained from Eq. (2) with the stochastic matrix (79) when its size NxN tends to infinity, i.e., $N \to \infty$. However, for a finite N, the solution to Eq. (2) is periodic. For instance, if N=4:

$$p_{jk}^{(n)} = \left[1 + (-1)^{k+j-n}\right], \qquad n = 1,2,\ldots etc. \tag{81}$$

i.e., the solution has period 2, and the process is not ergodic.

However, the finite-state Markov chains have the following property (even if the conditions (5), (36) or (41) do not hold):

$$\lim_{n\to\infty}\left\{\frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)}\right\} = \tilde{\pi}_j^{\infty} \tag{82}$$

where $\tilde{\pi}_i^{\infty}$ satisfy equations similar to (9), (37), or (40).

In our case

$$\tilde{\pi}_i^{\infty} = \sum p_{ij} \tilde{\pi}_j^{\infty} \tag{83}$$

Therefore, the expected average of the probability distribution $\tilde{\pi}^{\infty}$ approximates the ergodic solution (80), and it can be obtained from only one run of QRN.

5. Simulation of Chaos and Turbulence

In this section we will apply effective simulations of the Fokker-Planck equation by QRN discussed in the previous section, for solving the hardest problem of theoretical physics: prediction of chaotic and turbulent motions.

5.1 Background.

Many physical phenomena cannot be effectively predicted because their mathematical models are unstable in a sense that small errors in the initial conditions or in the forces grow exponentially. If the underlying dynamical system does not have an alternative stable state, this process continues indefinitely (chaos, turbulence), and computations performed by any classical computer fail. In order to illustrate that, suppose that the error ε characterizing the chaotic evolution, on average, grows exponentially:

$$\varepsilon = \varepsilon_0 e^{\gamma}, \quad \gamma = const > 0$$

If the accuracy to which the initial condition ε_0 can be presented is given by L, then during the time period

$$\Delta t < \frac{1}{\gamma} \ell n \frac{L}{\varepsilon_0}$$

two different trajectories for which

$$|\varepsilon_0' - \varepsilon_0''| < L$$

cannot be distinguished.

However, for $t > \Delta t$

these two trajectories diverge such that they must be considered as two different trajectories. Moreover, the distance between them tends to infinity as $t \to \infty$ no matter how small these distances were initially. This is why the evolution ones computed cannot be reproduced again. Hence, the confident time interval Δt^* during which the evolution can be predicted to accuracy of ε^* is:

$$\Delta t < \frac{1}{\gamma} \ell n \frac{\varepsilon^*}{L}$$

i.e., it grows logarithmically slow with increase of resolution.

The computational intractability of chaos can be elucidated by the argument provided by Wolfram[1]. He notes that a universal (classical!) computer can always simulate step by step the behavior of any physical system, but for stable systems it has certain shortcuts which allows one to predict events earlier than they actually happen; however, since there are no such shortcuts for chaotic systems, they, in principle,` cannot be predicted. In support of this argument, Ford [2] noted that for stable systems

$$t_c \propto \ell og_2 \, t_p$$

while for chaotic systems:

$$t_c \propto t_p$$

where t_c and t_p are the computational and physical time, respectively.

Thus, this brief review demonstrates that there is a strong belief among physicists and computer scientists that although chaos is theoretically computable (in a sense that it does not belong to the class of uncomputable problems), it is nevertheless intractable by

classical universal computers, and therefore, chaos attains the same computational status as those of NP-complete problems.

We will start with some comments to this statement. Despite the indisputable similarity between chaos and NP-complete problems from the viewpoint of computational complexity, there is a fundamental difference between them: chaos is a natural phenomenon while NP-complete problems are man-made, i.e., they can be mapped into decision problems with yes/no solutions. That is why the origin of the computational complexity of chaos is in the dynamical instability of the corresponding mathematical model. Hence, having in mind that infinite precision can never be achieved, the only way to predict chaos is to suppress this instability. One possible way to do that was introduced by the author in his earlier works [3-5]. In this paper, in view of potential availability of a quantum computer, a more effective way to perform the computational stabilization is proposed.

5.2 Governing Equations

Consider a system of nonlinear ODE:

$$\dot{v}_i = f_i(\{v\}), \qquad \{v\} = v_1, v_2, \ldots v_i \tag{84}$$

which is exponentially sensitive to errors in initial conditions. Then its solution eventually attains some stochastic-like properties, and it is representable in the Langevin form:

$$\dot{v}_i = f_i(\{v\}) + g_{ij}(\{v\}) v_j \Gamma_j(t) \tag{85}$$

where now v_i is the ensemble averaged value of the original velocities v_i, $\Gamma_i(t)$ is the Langevin force:

$$\overline{\Gamma}_i(t) = 0, \quad \overline{\Gamma}_i(t)\overline{\Gamma}_i(t) = 2\delta_{ij}\delta(t - t') \tag{86}$$

and the functions $g_{ij}(\{v\})$ are to be found.

Then a probabilistic description of the behavior of the solution $v_i = v_i(t)$ is expressed by the corresponding Fokker-Planck equation (see Eq. (70))

$$\frac{\partial \pi}{\partial t} = \left[-\sum_{i=1}^{'} \frac{\partial}{\partial v_i} D^{(1)}(\{v\}) + \sum_{i,j=1}^{'} \frac{\partial^2}{\partial v_i \partial v_j} D_{ij}^{(2)}(\{v\}) \right] \pi, \tag{87}$$

where $\pi(\{v\},t)$ is the $\{v\}$-distribution function, and

$$D_i^{(1)} = f_i(\{v\}) + g_{kj}(\{v\}) \frac{\partial}{\partial v_k} g_{ij}(\{v\})$$

$$D_{ij}^{(2)} = g_{ik}(\{v\}) g_{jk}(\{v\}) \tag{88}$$

In classical stochastic ODE, the functions g_{ij} are found from the statistical properties of the random external force. In our case such a force does not exist: it is the internal force generated by the mechanism of instability which replaces them. Consequently, the sought functions g_{ij} should be connected with the rate of instability of the original system (84). Since formally these functions can be interpreted as the Reynolds stresses (introduced in theory of turbulent motions), the problem of finding g_{ij} is actually equivalent to the closure problem in theory of turbulence, and it will be discussed below.

5.3 The Closure Problem.

The closure problem arose more than a hundred years ago after O. Reynolds introduced his famous decomposition of velocity field into mean and fluctuating components in the Novier-Stokes equations:

$$\rho\left(\frac{\partial \hat{v}}{\partial t}+\hat{v}\nabla\hat{v}\right)=-\nabla p+\nu\nabla^2\hat{v}, \qquad \nabla\hat{v}=0 \tag{89}$$

where $\hat{v}, \rho, p,$ and ν are velocity, density, pressure and viscosity, respectively.

After replacing spatial derivatives by the corresponding differences, . (89) takes the form of Eq. (84), and then the so called Reynolds stresses are represented by the functions g_{ij} in Eq. (85). In terms of Eq. (85), the closure problem is to find g_{ij} as functions of the averaged variables v_k, i.e.,

$$q_{ij}=q_{ij}(\{v\}),\{v\}=v_{1,\ldots}v_{l} \tag{90}$$

All earlier attempts to find the closure (90) based upon rhelogical or empirical considerations failed since they were problem specific. Our approach will be based upon the stabilization principle [11] which briefly can be described as follows below.

Because of finite resolution in computations and measurements, there are always such domains where different values of a variable cannot be distinguished, and as a mathematical idealization, the variable in these domains can be considered as having small random fluctuations of order inversely proportional to this resolution. In particular, such random fluctuations are present in the initial conditions of the system (84). When this system is stable, the fluctuations do not grow, and therefore, they can be ignored. However, if the system is unstable and there is no alternative stable state, the fluctuations grow until they start interacting with the mean values of the variables. Such interaction acting as a feedback eventually stabilizes the system on the level of marginal stability since

then the mechanism for further changes in fluctuations is eliminated. Mathematically this can be interpreted as follows: the original system (84) is unstable in the class of smooth functions, but it is stable in the class of multivalues (or random) functions. This interpretation is based upon the fact that the concept of stability is defined only subject to a certain class of functions in the same way in which convergence of a sequence is defined by the definition of the distance: it may converge in one space and diverge in another. Consequently, by introducing the stochastic force (or the Reynolds stresses) one enlarges the class of functions in which a stable solution is sought, and from this condition the closure (90) should be found.

Let us turn now to Eq. (84). Its local instability is caused by those eigenvalues of the matrix:

$$\left|\frac{\partial f_i}{\partial v_j}\right|_{\{v\}=\{v_0\}} = a_{ik}\left[\lambda_1,\ldots\lambda_n\right]a_{kj}^{-1}, \quad a_{ik}a_{kj}^{-1} = \delta_{ij} \tag{91}$$

which has positive real parts:

$$\mathrm{Re}\,\lambda_i^* > 0 \tag{92}$$

If the condition (92) is true for any state $\{v_0\}$ where the motion can occur, then the solution becomes chaotic. (Actually, this condition is too strong: it is sufficient, but not necessary; a sufficient and necessary condition is expressed in terms of Lyapunov exponents characterizing global stability).

Based upon the properties of the local stability defined by (91) and (92), one can now introduce a stabilizing force g_{ij} as follows:

$$g_{ij}(\{v\}) = a_{ik}\left[\lambda_1'\ldots\lambda_n'\right]a_{kj}^{-1} \tag{93}$$

where

$$\lambda_i' = \begin{cases} -\mathrm{Re}\,\lambda_1 & \text{if } \mathrm{Re}\,\lambda_1 > 0 \\ 0 & \text{otherwise} \end{cases} \tag{94}$$

By substituting Eq. (93) and (94) into Eq. (85), one can verify that now Eq. (85) is marginally stable in a small neighborhood of the state $\{x_0\}$.

5.4 Computational strategy.

For better physical interpretation, we will illustrate the computational strategy based upon the Novier-Stokes equation (89) in the discretized form (84) simultaneously with quantum simulations of the corresponding Fokker-Planck equation (87).

Suppose that the stationary version of Eq. (89) $(\text{with } \partial v/\partial t = 0)$is solved subject to some boundary conditions, and the corresponding laminar flow is obtained

$$v_i = v_i^0 \tag{95}$$

This means that all the parameters in Eqs. (93) and (94) are known, and the stochastic force in Eq. (88) can be found. If this force is zero, then the laminar motion is stable, and no further actions are needed. If this force is not zero,

$$g_{ij}^{(1)} \neq 0 \tag{96}$$

then the laminar flow is unstable, its evolution attains stochastic properties, and after substituting the force (96) in the Fokker-Planck equation (87), with the initial conditions

$\pi(t=0)=\delta(\{v\}-\{v\}_0)$, one obtains the stationary probability distribution of velocities $\{v\}$ which characterizes the new (turbulent) state:

$$\pi=\pi_{\infty}^{(1)}(\{v\}) \tag{97}$$

Next one can find the averaged (or expected) velocity field:

$$\overline{v}_i^{(1)}=\int_V v_i\pi_{\infty}^{(1)}(\{v\})dV \tag{98}$$

In general, these expected values may be different from the original laminar values in Eq. (95), i.e.,

$$v_i^{(1)}\neq v_i^{(0)} \tag{99}$$

Then, the stochastic force (99) should be recalculated based upon Eqs. (93) and (94):

$$g_{ij}=g_{ij}^{(2)} \tag{100}$$

and substituted back into Eq. (87) with the initial conditions (97). This iterative process has to be repeated until

$$\overline{v}_i^{(n-1)}=\overline{v}_i^{(n)} \tag{101}$$

Now one can appreciate the exponential speedup provided by QRN. indeed, the number of Eq. (91) in case of turbulent motions has the order of $10^5\div10^6$. This means that the Fokker-Planck equation (87) has the same number of independent variables, i.e.,

$$\{v\} = v_1, v_2 \cdots v_\ell, \qquad \ell \sim 10^6 \tag{102}$$

As pointed out in the previous section, the computational resources grow exponentially in $\ell\left(\sim\left(\frac{1}{\varepsilon}\right)^{\ell}\right)$, and that makes any classical approach unthinkable. At the same time, simulations by QRN requires resources which are independent upon $\ell\left(\sim\left(\frac{1}{\varepsilon^2}\right)\right)$. It should be recalled that the direct product representation (45) of the unitary matrix contracts the size of the input vector from N^n to mn, while the number of measurements $n\ell$ per unit time τ grows linearly in ℓ.

In addition to that, the size of the output which strictly speaking, is also exponentially large, can be contracted to polynomial size by selecting a polynomial number of critical points.

$$\{v^*\} = v_{i_1}, v_{i_2}, \ldots v_{i_\ell}, \quad i = 1,2,\ldots k \propto \ell \tag{103}$$

where the solution can deliver practically sufficient information about the properties of the velocity distribution such as expected velocity field, its variance, higher moments etc., and disregarding the solution at the rest of the points in Eq. (102).

It should be noticed that the speed of convergence of the iterations to the condition (101) is problem specific, and at this stage, we are not ready to make any comments on that.

5.5 Prediction of Chaos.

In order to further elucidate the connection between the Fokker-Planck equation and prediction of chaos, we will discuss a relatively simple example introduced first in [12]. Let us consider an inertial motion of a particle M of unit mass on a smooth pseudosphere S having a constant negative curvature (Fig. 7)

$$G_0 = \text{Const} < 0 \tag{104}$$

Remembering that trajectories of inertial motions must be geodesics of S, we will compare two different trajectories assuming that initially they are parallel and that the distance between them, ε_0, is very small.

As shown in differential geometry, the distance between such geodesics will exponentially increase:

$$\varepsilon = \varepsilon_0 e^{\sqrt{-G_0}\ t}, \quad G_0 < 0 \tag{105}$$

Hence, no matter how small the initial distance ε_0, the current distance ε tends to infinity.

Let us assume now that the accuracy to which the initial conditions are known is characterized by L. It means that any two trajectories cannot be distinguished if the distance between them is less than L, i.e., if:

$$\varepsilon < L \tag{106}$$

The period during which inequality (106) holds has the order:

$$\Delta t \sim \frac{1}{\left|\sqrt{-G_0}\right|}\ \ell n \frac{L}{\varepsilon_0} \tag{107}$$

However, for

$$t >> \Delta t \tag{108}$$

these two trajectories diverge such that they can be distinguished and must be considered as two different trajectories. Moreover, the distance between them tends to infinity even if ε_0 is small (but not infinitesimal). This is why the motion, once recorded, cannot be reproduced again (unless the initial conditions are known exactly), and consequently, it attains stochastic features. The Liapunov exponent for this motion is positive:

$$\sigma = \lim_{t \to \infty,\ d(0) \to 0} \frac{1}{t}\ \ell n\ \frac{\varepsilon_0 e^{\sqrt{-G_0}\ t}}{\varepsilon_0} = \sqrt{-G_0} = \text{Const} > 0 \qquad (109)$$

Let us introduce a system of coordinates at the surface S: the coordinate q_1 along the geodesic meridians, and the coordinate q_2 along the parallels. In differential geometry such a system is called semi-geodesical. The square of the distance between adjacent points on the pseudosphere is:

$$ds^2 = \overline{g}_{11}\ dq_1^2 + 2\overline{g}_{12} dq_1 dq_2 + \overline{g}_{22} dq_2^2 \qquad (110)$$

where

$$\overline{g}_{11} = 1,\ \overline{g}_{12} = 0,\ \overline{g}_{22} = -\frac{1}{G_0} e^{-2\sqrt{-G_0}\ q_1} \qquad (111)$$

The Lagrangian for the inertial motion of the particle M on the pseudosphere is expressed via the coordinates and their temporal derivatives as:

$$L = \overline{g}_{ij} \dot{q}_i \dot{q}_j = \dot{q}_1^2 - \frac{1}{G_0} e^{-2\sqrt{-G_0}\ q_1}\ \dot{q}_2^2 \qquad (112)$$

and consequently

$$\frac{\partial L}{\partial q_2} = 0 \tag{113}$$

while

$$\frac{\partial L}{\partial q_1} \neq 0, \quad if \quad \dot{q}_2 \neq 0 \tag{114}$$

Hence, q_1 and q_2 play roles of position and ignorable coordinates, respectively. Therefore, an inertial motion of a particle on a pseudosphere is stable with respect to the position coordinate q_1, but it is unstable with respect to the ignorable coordinate. It can be shown that such a motion becomes stochastic [12].

The governing differential equations follow from the Lagrangian (110):

$$\dot{v}_1 = -\frac{1}{\sqrt{-G_0}} e^{-2\sqrt{-G_0}\ q_1}\ v_2^2, \quad \dot{q}_1 = v_1$$

$$\dot{v}_2 = -2\sqrt{-G_0}\ \ v_1 v_2, \qquad \dot{q}_2 = v_2 \tag{115}$$

and, as shown above, their solution is chaotic. It can be verified [11] that the stabilizing stochastic force (93), (94) which suppresses the Lyapunov exponent (109) down to zero must have the potential energy:

$$\prod = -EG_0 q_1^2, \quad G_0 < 0, \tag{116}$$

where E is the kinetic energy of the particle. Substituting (116) in the Lagrangian (110) one rewrites Eqs. (115) in the Langevin form:

$$\dot{v}_1 = \frac{1}{\sqrt{-G_0}} e^{-2\sqrt{-G_0}\,\bar{q}_1}\ \bar{v}_2^2, \quad \dot{\bar{q}}_1 = \bar{v}_1 \tag{117}$$

$$\dot{\bar{v}}_2 = -2\sqrt{-G_0}\ \bar{v}_1 \bar{v}_2 - EG_0\,\bar{q}_1^2\ \ \Gamma(t), \quad \dot{\bar{q}}_2 = \bar{v}_2$$

and, based upon Eqs. (87), and (88), one obtains the corresponding Fokker-Planck equation:

$$\frac{\partial \pi}{\partial t} = \alpha\pi + \beta_1 \frac{\partial \pi}{\partial v_1} + \beta_2 \frac{\partial \pi}{\partial v_2} + \beta_3 \frac{\partial \pi}{\partial q_1} + \beta_4 \frac{\partial \pi}{\partial q_2} + E^2 G_0^2 \frac{\partial^2 \pi}{\partial v_2^2} \cdots \qquad (118)$$

where $$\alpha = \frac{2}{\sqrt{-G_0}} v_1, \beta_1 = -\frac{1}{\sqrt{-G_0}} e^{-2\sqrt{-G_0}\, q_1} \quad \beta_3 = v_1, \beta_4 = v_2$$

Thus, the probability distribution $\pi(v_1, v_2, q_1, q_2, t)$ is described by Eq. (118) which is supposed to be solved subject to the following initial and boundary conditions, respectively:

$$\pi|_{t=0} \quad \delta[v_1 - v_1(t=0),\ v_2 - v_2(t=0),\ q_1 - q_1(t=0) q_2 - q_2(t=0)] \qquad (119)$$

$$\int_V \pi \, dv_1 \, dv_1 \, dq_1 \, dq_2 = 1$$

It should be noticed that the solution to Eq. (118) represents the exact description of an inertial motion of a particle on a pseudosphere; this motion is stochastic, while the stochasticity is generated by the orbital (chaotic) instability. The closed form expression (116) for the stochastic force is due to a remarkable property of the system (115): it has constant exponential divergence of the trajectories (105) which simply defines the positive Lyapunov exponent, (see Eq. 105). That is why in this case there are no iterative recalculations of the stochastic force needed.

Although Eq. (118) is relatively simple, its closed form solution is not available, i.e., numerical computations or quantum simulations must be applied.

6. Information Processing

6.1 Data Compression

Most data to be processed are not totally random: they are correlated. Because of that, they contain some amount of redundancy which can be removed when data are stored,

and replaced when they are reconstructed. Elimination of such a redundancy can be associated with data compression.

Data describing natural phenomena (fields of velocities, temperatures, forces, images) most likely have a tensor structure. But not all of the tensor components are equally important: some of them, called invariants, describe the fundamental properties of physical objects, and they deliver the largest portion of useful information; other components which depend upon object orientations in space are less informative. In many cases a tensor can be composed (exactly or approximately) into a direct product of tensors of lower dimensionality, and that leads to significant reduction of the number of informative parameters. The effectiveness of QRN for data compression can be associated with the last case. Indeed, as shown above (see Eq. (69)), an exponentially large number S components of the tensor (40) can be generated based upon much lower number q of independent parameters. Hence, formally, QRN can effectively compress such data which have high Shannon complexity

$$S = 2^{n'} \tag{120}$$

and low algorithmic complexity:

$$q = 4n \tag{121}$$

with the compression ratio

$$\xi = \frac{2^{n'}}{4n} \tag{122}$$

However, the hardest part of the task is to solve the inverse problem: given N' data (40), find the unitary matrix in the form of (44), or (48) and the interference vector (32) which

generate a stochastic attractor whose tensor components are equal (or close) to those of (40).

Analytically the problem can be formulated as follows: by an appropriate choice of QRN parameters, minimize the function:

$$E = \sum_{i_k=1}^{N} \left(\pi_{i_1, i_2 \cdots i_l} - \pi^*_{i_1, i_2 \cdots i_l} \right)^2 \tag{123}$$

Here $\pi^*_{i_1 \cdots i_l}$ represents a datum out of N' given data, and $\pi^*_{i_1 \cdots i_l}$ are the components of the sought ergodic attractor whose components found from Eq. (40) should match given data.

The QRN parameters to be optimized can be represented by components of the unitary matrix (48), or by the components of the interference vector (28). For the sake of concreteness, we will choose the latter. Assuming that the interference vector is represented as a direct and dot products of two component vectors (see Eq. (46) and (48) one has to vary in Eq. (123) q'independent parameters where

$$q' = 3nm \tag{124}$$

However, the N' variables $\pi_{i_1 \ldots i_l}$ in (Eq. 123) depend upon the interference vector only via the components of the stochastic matrix $p_{12 \ldots l}$ (see Eqs (30) and (33), i.e., a system of N' equations (40) should be solved prior to minimization (123).

Hence a gradient descent approach to minimization of the function (123) which represents a typical way of learning in classical neural nets, in this particular case would require exponential computational resources. So one arrives exactly at the same situation which was discussed in the two previous sections.

As an alternative approach, one can try a learning by QRN simulations based upon random gradient descent. That is how it can be done.

First, select (randomly) an initial interference vector $a_0\left(|a_0|^2 = 1\right)$ run QRN and find E_0 according to Eq. (123). If $E_0 = 0$, the process is finished. If $E_0 > 0$, generate randomly a new interference vector $a_0 + \Delta a\left(|a_0 + \Delta a| = 1\right)$ where $|\Delta a| << 1$, run QRN again and find E_1. If $E_1 < E_0$, continue changing a in the same "direction," if $E_1 > E_0$ then switch the "direction" to the opposite one. If, after several runs, the difference $E_n - E_{n-1}$ changes its sign, stop the process and initiate randomly new "direction," etc.

Eventually the process will end with $E_\infty = 0$, or $E_\infty > 0$,. In the first case the data compression is lossless, in the second case the loss is characterized by the value of E_∞.

Clearly, the direction of descent here is selected randomly, and therefore, it is not optimal. However, the procedure does not require implementation of an exponential number of constraints (see Eqs. (70) and (74)): they are automatically implemented by QRN, and that makes this quantum learning alternative more attractive. It should be noticed, that, strictly speaking, exact evaluation of E in Eq. (123) requires exponential resources. But QRN allows one to select a polynomial subset of components in Eq. (123) and evaluate only them in the same way in which it was described in the previous section. Similar simplification cannot be done in case of conventional gradient descent approach since all the Eqs. (37) are coupled.

Thus QRN can perform data compression in the following way: the N' data represented (exactly or approximately) as a component of a tensor which, in turn, is parametrized by much smaller number of independent components of the interference vector a. This vector stores all the information about the original data: being introduced to the QRN, it reconstructs them. Obviously one can store or transport the interference vector a which is composed of nm particles (see Eq. (46) and (48) in a classical, or if possible, in a quantum way.

6.2 Recognition of Pattern Combinations.

It would be, probably, unwise to exploit QRN for the same type of tasks (associative memory, pattern recognition) for which classical neural nets are utilized, and the reason for that is the following: QRN are linear in the probabilistic space, which means that all different initial vectors introduced to it eventually converge to the same stochastic attractor. In other words, QRN do not have an ability to discriminate. On the other hand, if an initial vector is introduced at each iteration in the form of the interference vector (see Eq. (32)), then one arrives at another extreme: each initial vector converges to its own stochastic attractor. But that means that several different vectors introduced to QRN simultaneously (in the form of a normalized vector sum) will converge to a new attractor which is different from each of the individual attractors. Physically this is a result of quantum interference between different initial vectors, and such an effect does not exist in classical neural nets.

In order to elucidate the importance of that, suppose that each initial vector is identified with a letter out of a certain collection of letters forming an alphabet. If introduced separately, each letter has its own image as a stochastic attractor in the probabilistic space of QRN. But if a group of letters is introduced simultaneously, their image in the probabilistic space of QRN can be interpreted as a word which has its own meaning. And that meaning can be very rich since the attractor has a very high Shannon complexity which is characterized by N' tiny features. In the same way, several words form sentences, etc. The rules, or grammar, which implement such a logical structure, are uniquely defined by the components U_{ij} of the unitary matrix (45): each collection of these components correspond to a different language. It is important to emphasize the fact that all the images of letters, words and sentences are represented by dynamical attractors: any small distortions of intitial vectors are suppressed by the contraction properties of the

process of attraction since the "distance" between two different initial vectors eventually becomes smaller and smaller.

To conclude this very brief description of the formation of new meanings by QRN, we will make some comments concerning possible philosophical interpretation of this phenomenon.

Indeed, it was always difficult to understand how biological neural nets can learn patterns of external world without any preliminary structure built-in to their synaptic interconnections. The experience with artificial neural nets shows that training without a preliminary structure is exponentially longer than those with a structure, and that poses the following question: who created the "first" structure in biological neural nets which provides the ability to learn and select useful properties in polynomial time? In other words, can natural selection act without a "creator"? The quantum neural nets may give a positive answer to this question: the logical structure of synaptic interconnections can be imposed by natural laws of physics, and in particular, by quantum mechanics. Hence, if biological neural nets utilize quantum effects in their performance, they may be able to learn the model of the external world, including its logical structure, in polynomial time without any preliminary structure.

7.0 Generating Stochastic Processes.

As shown above, QRN can be viewed as a universal and compact generator of stochastic processes, that cannot be achieved, even in principle, by any classical device. Indeed, it can generate multivariate Markov and non-Markov, linear and nonlinear stochastic processes with prescribed properties by simply changing a quantum interference pattern (see Eq. (32), (33) or the unitary matrix itself.

One of the most important application of simulated stochastic processes is the Monte-Carlo methods discussed in the Introduction. It provides approximate solutions

to a variety of mathematical problems such as solutions to algebraic, differential and integral equations, as well as to combinatorial problems. In particular, the Monte-Carlo approach appears to be effective in approximate solutions to enumeration problems (perfect matching in a bipartite graph, estimation of matrix permanent, estimation of the volume of a convex n-dimensional body, etc.) forming a class of #p-complete problems. (The #p-complete problems are known to be at least as hard as NP-complete problems). The fastest known algorithm for exact computations of the permanent requires $0(n2^n)$ operations. In case of (ε,δ) -approximation (where the algorithm produces an approximation of the permanent with relative error less than ε with the probability greater than $1-\delta$) the required time is still exponential, $0(2^{n/2})$. As shown in [13] , by an appropriate selection of the Markov chain, the same problem in (ε,δ)-approximation can be solved in polynomial time. However, the main restriction to application of the Monte-Carlo approach is generation of true stochastic processes (the rapidly generating algorithms in common use produce sequences whose properties depart from this ideal with error proportional to n). That is why QRN can significantly amplify the effectiveness of the Monte-Carlo methods.

The second area of application is performance of sampling experiments on the model of the systems. In this area not only the limit probability distributions, but their time evolutions are important as well. And in this connection, the nonlinear stochastic processes which allows one to control the current strategy in real time by changing the stochastic attractors and concentrating probabilities in a certain domain (depending upon a changing objective) become very useful in modeling decision making process in a game-type situation.

In this section we will briefly discuss the strategy for simulating stationary stochastic processes by QRN. We will start with the following problem: find a unitary matrix U and an interference vector $|\psi'\rangle$ (see Eq. (27) such that the corresponding

stochastic process converges to a stationary attractor with prescribed probability distribution:

$$\pi^{\infty} = \left(\pi_1^{\infty}, \ldots \pi_N^{\infty}\right), \quad N = 2^n \tag{125}$$

Clearly the first step is to find an appropriate stochastic matrix P which leads to the distribution (125). As follows from Eqs. (9), there is an infinite number of such matrices, but only one of them provides the shortest transient time to the attractor. In order to find this matrix, one should recall that (ε, δ) -approximation can be solved in polynomial time. However, the main restriction to application of the Monte-Carlo approach is generation of true stochastic processes (the rapidly generating algorithms in common use produce sequences whose properties depart from this ideal with error proportional to n)

$$P^{\tilde{t}} = \begin{pmatrix} \pi_1^{\infty} \pi_2^{\infty} \ldots \pi_N^{\infty} \\ \vdots \ \vdots \quad \vdots \\ \pi_1^{\infty} \pi_2^{\infty} \ldots \pi_N^{\infty} \end{pmatrix} \tag{126}$$

if $\tilde{t}$ is the discrete time (i.e., the number of steps) to the attractor.

It is easily verifiable that the matrix (126) satisfies Eqs. (9), and, at the same time, it completely eliminates the transition period to the attractor. Hence, now in Eqs. (125) and (126) the super-index ∞ can be omitted:

$$\pi = \left(\pi_1 \ldots \pi_N\right), \quad P = \begin{pmatrix} \pi_1 & \cdots & \pi_N \\ \cdots & \cdots & \cdots \\ \pi_1 & \cdots & \pi_N \end{pmatrix}, \quad N = 2^n. \tag{127}$$

Obviously

$$P=\begin{pmatrix}1 & \cdots & 1\\ \cdots & \cdots & \cdots\\ 1 & \cdots & 1\end{pmatrix}\begin{pmatrix}\pi_1 & & 0\\ & \ddots & \\ 0 & & \pi_N\end{pmatrix} \tag{128}$$

while

$$\begin{pmatrix}1 & \cdots & 1\\ \cdots & \cdots & \cdots\\ 1 & \cdots & 1\end{pmatrix}=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\otimes\ldots\otimes\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix} \tag{129}$$

The diagonal matrix in Eq. (128) can be approximated by a direct product of 2x2 diagonal matrices:

$$\begin{pmatrix}\pi_1 & & 0\\ & \ddots & \\ 0 & & \pi_N\end{pmatrix}\cong\begin{pmatrix}\pi_1^0 & 0\\ 0 & 1-\pi_1^0\end{pmatrix}\otimes\cdots\otimes\begin{pmatrix}\pi_n^0 & 0\\ 0 & 1-\pi_n^0\end{pmatrix} \tag{130}$$

under the condition that $\overset{o}{\pi}_1,\ldots\overset{o}{\pi}_n$ minimize the sum:

$$\left(\pi_1-\overset{o}{\pi}_1\overset{o}{\pi}_2\ldots\overset{o}{\pi}_n\right)^2+\ldots+\left[\pi_N-\left(1-\overset{o}{\pi}_1\right)\left(1-\overset{o}{\pi}_2\right)\ldots\left(1-\overset{o}{\pi}_n\right)\right]^2\rightarrow\min \tag{131}$$

Therefore,

$$P\cong\begin{pmatrix}\overset{o}{\pi}_1 & 1-\overset{o}{\pi}_1\\ \overset{o}{\pi}_1 & 1-\overset{o}{\pi}_1\end{pmatrix}\otimes\cdots\otimes\begin{pmatrix}\overset{o}{\pi}_n & 1-\overset{o}{\pi}_n\\ \overset{o}{\pi}_n & 1-\overset{o}{\pi}_n\end{pmatrix} \tag{132}$$

Consequently, the corresponding unitary matrix U as well as the interference vector $|\varphi'\rangle$ can be sought in a similar form

$$U = U_1 \otimes \ldots \otimes U_n,\ U_j = \begin{pmatrix} u_{11}^{(j)} & u_{12}^{(j)} \\ u_{21}^{(j)} & u_{22}^{(j)} \end{pmatrix} j, \quad j = 1,2,\ldots n \tag{133}$$

$$|\psi'\rangle = |\psi'_{(n)}\rangle \otimes \ldots |\psi'_{(n)}\rangle, |\psi'_{(i)}\rangle = \begin{pmatrix} a_1^{(i)} \\ a_2^{(i)} \end{pmatrix}, i = 1,2,\ldots n \tag{134}$$

Now the problem of finding U and $|\psi'\rangle$ is reduced to finding the two dimensional components in Eqs. (133) and (134) based upon the corresponding two-dimensional components in Eq. (132).

If a unitary matrix U_j and an interference vector $|\psi'_{(\gamma)}\rangle$ are sought in the form

$$\begin{aligned} u_{11}^{(j)} &= e^{\frac{i\alpha^{(j)}}{2} + \frac{i\beta^{(j)}}{2} + i\delta^{(j)}} Cos\left(\frac{\theta^{(j)}}{2}\right) \\ u_{12}^{(j)} &= e^{\frac{i\alpha^{(j)}}{2} - \frac{i\beta^{(j)}}{2} + i\delta^{(j)}} Sin\left(\frac{\theta^{(j)}}{2}\right) \\ u_{21}^{(j)} &= -e^{\frac{-i\alpha^{(j)}}{2} + \frac{i\beta^{(j)}}{2} + i\delta^{(j)}} Sin\left(\frac{\theta^{(j)}}{2}\right) \\ u_{22}^{(j)} &= e^{\frac{-i\alpha^{(j)}}{2} - \frac{i\beta^{(j)}}{2} + i\delta^{(j)}} Cos\left(\frac{\theta^{(j)}}{2}\right) \end{aligned} \tag{135}$$

$$a_1^{(j)} = \overset{o(j)}{a_1}\, e^{i\gamma_1^{(j)}}, \qquad a_2^{(j)} = \overset{o(j)}{a_2}\, e^{i\gamma_2^{(j)}} \tag{136}$$

then the eight parameters

$$\alpha^{(j)}, \beta^{(j)}, \delta^{(j)}, \theta^{(j)}, \overset{o(j)}{a_1}\ \overset{o(j)}{a_2}\ \gamma_1^{(j)}, \text{ and } \gamma_2^{(j)} \tag{137}$$

are supposed to satisfy the following four equations:

$$\frac{\left|u_{11}^{(j)}\left(a_1^{(j)}+1\right)+u_{12}^{(j)}a_2^{(j)}\right|^2}{\left|a_1^{(j)}+1\right|^2+\left|a_2^{j}\right|^2}=\overset{o}{\pi}_j=\frac{\left|u_{11}^{(j)}a_1^{(j)}+u_{12}\left(a_2^{(j)}+1\right)\right|^2}{\left|a_1^{(j)}\right|^2+\left|a_2^{(j)}+1\right|^2}\cdots \tag{138}$$

$$\frac{\left|u_{21}^{(j)}\left(a_1^{(j)}+1\right)+u_{22}^{(j)}a_2^{(j)}\right|^2}{\left|a_1^{(j)}+1\right|^2+\left|a_2^{(j)}\right|^2}=1-\overset{o}{\pi}=\frac{\left|u_{21}^{(j)}a_1^{(j)}+u_{22}^{(j)}\left(a_2^{(j)}+1\right)\right|^2}{\left|a_1^{(j)}\right|^2+\left|a_2^{(j)}+1\right|^2}\cdots \tag{139}$$

For $$\overset{o}{\pi}_j=\frac{1}{2} \tag{140}$$

i.e., for a uniform distribution, the system (138), (139) has a trivial solution

$$\theta^{(j)}=\frac{\pi}{2},\ \alpha^{(j)}=\beta^{(j)}=\delta^{(j)}=\gamma_1^{(j)}=\gamma_2^{(j)}=0;\ \overset{o(j)}{a_1}=\overset{o(j)}{a_2}=0 \tag{141}$$

In general, all the parameters in (137) will be functions of the only one variable $\overset{o}{\pi}$; and these functions can be tabulated. Some freedom in choice of the parameters (137) can be exploited for the purpose of simple implementation.

Similar strategy can be applied for simulating multi-variate stochastic processes with utilization of Eqs. (25), (33) and (48) instead of Eqs. (3), (30), and (44), respectively

In conclusion of this section, we will illustrate how a prescribed probability distribution or a stochastic matrix can be achieved by an appropriate choice of a unitary matrix U and an off-set vector $\bar{a}$.

Example 1

Given a double valued random variable:

$$p_r(z_i = 0) = p_r(z_i = 1) = \frac{1}{2} \tag{142}$$

Find the corresponding unitary matrix.

Solution:

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \rightarrow p = \begin{pmatrix} ½ & ½ \\ ½ & ½ \end{pmatrix} \tag{143}$$

The number of measurement per iteration

$$m = 2 \tag{144}$$

The number of different reset operations

$$r = 4 \tag{145}$$

The off-set vector is not needed.

The solution (142) is the simplest, and it eliminates a transition period to the attractor.

Example 2

Given a random variable ξ uniformly distributed over an interval of the length $\ell = 2^n$

$$p_r(k \le \xi \le k+1) = \frac{1}{\ell} \tag{146}$$

Find the corresponding unitary matrix.

Solution:

$$U = U_1 \otimes \ldots U_n, \quad U_k = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{147}$$

assuming that ξ has a binary representation:

$$\xi = z_1 \cdot 2^{-1} + z_2 \cdot 2^{-2} + \ldots + z_n \cdot 2^{-n} \tag{148}$$

where z_i are defined by Eq. (142)

The number of measurement per iteration

$$m = 2n \tag{149}$$

The number of different reset operations

$$r = 4n \tag{150}$$

Remark: if ξ is given on a hypercube $\{0,1\}^k$, then

$$U = U^{(1)} \otimes \ldots \otimes U^{(k)}, \quad U^{(i)} = U_1^{(i)} \otimes \ldots \otimes U_n^{(i)} \tag{151}$$

where $U_j^{(i)}$ has the form (147), while

$$m = 2nk, \quad r = 4nk \tag{152}$$

The off-set vector is not needed.

Example 3

Given a double valued random variable:

$$p_r(z_i = 0) = \pi, \quad p_r(z_i = 1) = 1 - \pi, \quad 0 < \pi < 1 \tag{153}$$

Find the corresponding unitary matrix U and an off-set vector $\bar{a}$

Solution:

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \bar{a} = \{a_1, a_2\} \tag{154}$$

$$a_1 = a_2 = \frac{1}{2}\left(-1 \pm \sqrt{1 + \frac{2\pi - 1}{1 - \pi}}\right) \tag{155}$$

$$m = 2, \quad r = 4 \tag{156}$$

Remark: Eqs. (154), (155) express U and $\bar{a}$ in the simplest form.

Indeed, assuming that

$$U=\begin{pmatrix} Cos\varphi & Sin\varphi \\ -Sin\varphi & Cos\varphi \end{pmatrix},\ a_1=a_2=a,\ I_m\, a=0 \tag{157}$$

one finds that $\varphi=\frac{1}{\sqrt{2}}$ from the equality:

$$\left[(a+1)Cos\varphi+a\,Sin\varphi\right]^2=\left[a\,Cos\varphi+(a+1)Sin\varphi\right]^2 \tag{158}$$

Then a is found from the condition:

$$\frac{\frac{1}{2}\left[(a+1)^2+a^2+2a(a+1)\right]}{(a+1)^2+a^2}=\pi \tag{159}$$

Example 4

Given a random variable ξ distributed over an interval of the length $\ell=2^n$. Assume that its probability can be defined by n independent parameters (moments, Fourier coefficients, etc). In particular, assume that ξ is sought in a binary form (148).

$$\xi=\sum_{j=1}^{n}z_j\cdot 2^{-j};\ p_r\left(z_j=0\right)=\pi,\ p_r\left(z_j=1\right)=1-\pi_j \tag{160}$$

and the probabilities of n (out of 2^n) particular configurations of ξ are prescribed as $\overset{o}{\pi}_j$, $j=1,2,\ldots n$. For instance, take first $q=\ell og_2\ n$ terms in Eq. (160) and prescribe probabilities for all n different configurations of these terms:

$$\underbrace{00\ldots0}_{q}\ [\text{the rest}]=\overset{o}{\pi}_1$$

$$\underbrace{00\ldots01}_{q}\ [\text{the rest}]=\overset{o}{\pi}_2 \tag{161}$$

$$\underbrace{11\ldots1}_{q}\ [\text{the rest}]=\overset{o}{\pi}_n$$

Then

$$\pi_1\pi_2\cdots\pi_n=\overset{o}{\pi}_1$$

$$\pi_1\pi_2\cdots\pi_{n-1}(1-\pi_n)=\overset{o}{\pi}_2 \tag{162}$$

$$(1-\pi_1)\cdots(1-\pi_n)=\overset{o}{\pi}_n$$

whence

$$\frac{\pi_n}{1-\pi_n}=,\frac{\overset{o}{\pi}_2}{\overset{o}{\pi}_1}\ \ i.e.,\qquad \pi_n=\frac{\overset{o}{\pi}_2/\overset{o}{\pi}_1}{1+\overset{o}{\pi}_2/\overset{o}{\pi}_1} \tag{163}$$

$$\frac{\pi_{n-1}}{1-\pi_{n-1}}=\frac{\overset{o}{\pi}_3}{\overset{o}{\pi}_2}, i.e., \qquad \pi_{n-1}=\frac{\overset{o}{\pi}_3/\overset{o}{\pi}_2}{1+\overset{o}{\pi}_3/\overset{o}{\pi}_2} \tag{164}$$

etc.

Eventually, all the probabilities $\pi_j(j=1,2\ldots n)$ are found in n steps given the distribution (161).

Hence, the corresponding U and $\bar{a}$ providing the distribution (161) can be based upon the examples 2 and 3.

Solution:

$$U=U_1\otimes\ldots U_n; \quad \bar{a}=\bar{a}_1\otimes\ldots\otimes\bar{a}_n \tag{165}$$

where

$$U_j=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}, \quad \bar{a}_j=\{a_j,a_j\} \tag{166}$$

and

$$a_j=\frac{1}{2}\left(-1\pm\sqrt{1+\frac{2\pi_j-1}{1-\pi_j}}\right) \tag{167}$$

here

$$m=2n, \quad r=4n \tag{168}$$

Remark 1. If ξ is given on a hypercube $\{0,1\}^k$, and all the k-components are independent, then Eqs. (10) and (11) can be applied.

Remark 2. If ξ is given on a hypercube $\{0,1\}^k$, and the k-components are correlated, in general one can find an equivalent one-dimensional distribution, and therefore, to utilize the example 4.

Example 5. given a 2x2 ("success" and "failure") stochastic matrix.

$$p = \begin{pmatrix} p_1 & 1-p_1 \\ p_2 & 1-p_2 \end{pmatrix}, \quad 0 < p_1 \neq p_2 < 1 \tag{169}$$

Find U and $\bar{a}$

Solution

$$\text{if } U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \textit{ and } \operatorname{Im}\bar{a} = 0 \tag{170}$$

then
$$\frac{(a_1+1)^2}{(a_1+1)^2 + a_2^2} = p_1; \qquad \frac{a_1^2}{a_1^2 + (a_2+1)^2} = p_2 \tag{171}$$

whence

$$a_1 = \frac{1+\sqrt{\frac{1}{p_1}-1}}{\sqrt{\frac{1}{p_2}-1} \quad -\sqrt{\frac{1}{p_1}-1}},$$

$$a_2 = \sqrt{\frac{1}{p_1}-1} \quad \frac{1+\sqrt{\frac{1}{p_2}-1}}{\sqrt{\frac{1}{p_2}-1} \quad -\sqrt{\frac{1}{p_1}-1}} \tag{172}$$

Example 6. Given an arbitrary nxn stochastic matrix

$$p = \begin{pmatrix} \overset{o}{p}_{11} & \cdots & \overset{o}{p}_{1n} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \overset{o}{p}_{n1} & \cdots & \overset{o}{p}_{nn} \end{pmatrix}, \qquad 0 < p_{ij} < 1 \tag{173}$$

Find U and $\bar{a}$

The matrix p performs mappings:

$$\overset{o}{\pi}\, p = \pi \tag{174}$$

where π is a probability vector.

First represent $\overset{o}{\pi}$ in a binary form (160). Then its transition to the vector π in Eq. (174) is equivalent to the transitions of each component in Eq. (160):

$$\left(\overset{o}{\pi}_i,\, 1-\overset{o}{\pi}_i\right)\begin{pmatrix} p_1^{(i)} & 1-p_1^{(i)} \\ p_2^{(i)} & 1-p_2^{(i)} \end{pmatrix} = \left(\pi_i, 1-\pi_i\right),\ i = 1,2\ldots n \tag{175}$$

As follows from Eq. (179), the probability of the transition

$$(10\ldots0) \rightarrow (10\ldots0) \tag{176}$$

is equal to p_{11} .

But as follows from Eqs. (160) and (175), the same transition has the probability $p_1^{(1)} p_1^{(2)} \ldots p_1^{(n)}$ hence,

$$\overset{o}{p}_{11} = p_1^{(1)} p_1^{(2)} \ldots p_1^{(n)} \tag{177}$$

Similarly,

$$\overset{o}{p}_{12} = \left(1 - p_1^{(2)}\right), \ldots p_1^{(n)} \tag{178}$$

etc.

Whence

$$\frac{p_1^{(1)}}{1 - p_1^{(1)}} = \frac{\overset{o}{p}_{11}}{\overset{o}{p}_{12}}, \quad i.e., \quad p_1^{(1)} = \frac{\overset{o}{p}_{11} \Big/ \overset{o}{p}_{12}}{1 + \overset{o}{p}_{11} \Big/ \overset{o}{p}_{12}} \tag{179}$$

Eventually, all the elements of the 2x2 stochastic matrices in Eqs. (175) can be found in n steps.

As in the Example 4, it is assumed here that the 2^{2n} components $\overset{o}{p}_{ij}$ in Eq. (173) can be derived from 2n independent parameters. (This is always the case when the matrix (173) results from a discretization of an underlying Fokker-Planck equation).

Then the solution to the problem is unique:

Solution:

$$U = U_1 \otimes \ldots \otimes U_n \qquad \bar{a} = \bar{a}^{(1)} \otimes \ldots \otimes \bar{a}^{(n)} \tag{180}$$

where

$$U_j = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \bar{a}^{(1)} = a_1^{(1)}, a_2^{(i)}, \quad Im\,\bar{a} = 0 \tag{181}$$

and

$$a_1^{(i)} = \frac{1 + \sqrt{\frac{1}{p_1^{(i)}} - 1}}{\sqrt{\frac{1}{p_2^{(i)}} - 1} - \sqrt{\frac{1}{p_1^{(i)}} - 1}}$$

$$a_2^{(i)} = \sqrt{\frac{1}{p_1^{(i)}} - 1} \quad \frac{1 + \sqrt{\frac{1}{p_2^{(i)}} - 1}}{\sqrt{\frac{1}{p_2^{(i)}} - 1} - \sqrt{\frac{1}{p_1^{(i)}} - 1}} \tag{182}$$

8.0 Conclusion

Thus, it has been demonstrated that quantum recurrent net as an analog device can be based upon a sequence of quantum and classical computations. during the quantum regime, a stochastic input pattern is transformed (according to Schrodinger equation) into the output stochastic pattern of the same dimensionality. During the following classical regime which includes quantum measurements and reset, the stochastic pattern is contracted into a pattern of lower dimensionality, and this contraction is equivalent to the performance of a sigmoid function. The combined effect of the alternating quantum and classical computations can be described by generalized random walk, i.e., by Markov chains in the

form of the Chapman-Kolmogorov equation. Eventually the output pattern approaches an attractor (which can be static, periodic, or ergodic), and such attractors can be utilized for storing certain patterns. The assignment of an appropriate unitary matrix can be based upon the optimal choice of the time period of the regime of quantum computations which actually represents the procedure of learning. But in addition to that, the transition probability matrix can be controlled by combining the output vector with an appropriately chosen interference vector.

Let us now summarize advantages of quantum neural nets.

The first fundamental advantage of QRN is based upon the property that all the computations are performed in a high-dimensional (abstract) space which is induced by quantum simulations of a physical device manipulating by vectors of much lower dimensionality, and that leads to exponential speedup and exponential capacity in QRN performance. These effects are amplified by the direct-product decomposability of the Hilbert space in which physical objects are defined based upon this property. Effective simulations of high-dimensional Fokker-Planck equation with applications to chaos, turbulence, combinatorial optimization and data compression were proposed and discussed.

The second advantage of QRN is in generation of true randomness which has been incorporated in a different kind of stochastic process of prescribed complexity. This property can be exploited for Monte-Carlo simulations, randomized algorithms, and simulations of complex patterns of behavior in physics, biology and social dynamics.

The third advantage of QRN is based upon interference between different patterns-inputs. Due to this interference the stored patterns acquire a logical structure in a sense that each combination of patterns has a qualitatively new meaning in the same way in which combinations of letters forming words do. This property opens up a way for storing retrieving and recognition of collections of patterns which may have new joint properties.

The problems of hardware implementations of quantum devices have not been discussed in this paper. However since the quantum nets operate by interleaving quantum

evolution with measurement and reset operations, they are far less sensitive to decoherence than other designs of quantum computers.

"The research described in this paper was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology and was sponsored by the National Aeronautics and Space Administration, Office of Space Science and Technology."

1. R. Feynman, Int. J. of Theoretical Physics, Vol. 21. No. 6/7, 1982.

2. J. Taub and H. Wozniakowski, "Breaking Intractability," Scientific American, January 1994, Vol. 46.

3. C. P. Williams and S. H. Clearwater, "Explorations in Quantum Computing," TELOS/Springer-Verlag, ISBN:038794768X (1997).

4. V. Cerny, Phys. Rev. A, Vol. 4B, No. 1, p 116, 1993.

5. M. Zak, Quantum Resonance for Simulations NP-complete Problems, Proceedings of NASA QCQC'98, Palm Springs, CA Feb. 16-20, 1998.

6. M. Zak, C. Williams, Quantum Neural Nets, Int. J of Theor. Physics, Feb. 1998.

7. M. Bartlett "An Introduction to Stochastic Processes," Cambridge University Press, 1966, p.33.

8. M. Schroeder, Fractals, Chaos, Power Laws, 1991, W. H. Freeman and Co.

9. Wolfram, S., Physics. Rev. Lett, 54, 735 (1985).

10. Ford, J. Quantum Chaos in "Directions in Chaos, Vol. 2. Edited by Hao-Bai-Lin, 1988.

11. M. Zak, etc., From Instability to Intelligence, Springer, 1997.

12. V. Arnold, Mathematical Methods of Classical Mechanics, Springer, Berlin, NY p. 331, 1988.

13. P. Motwani and P. Raghowan, Randomized Algorithms, Cambridge Univ. Press, 1995.

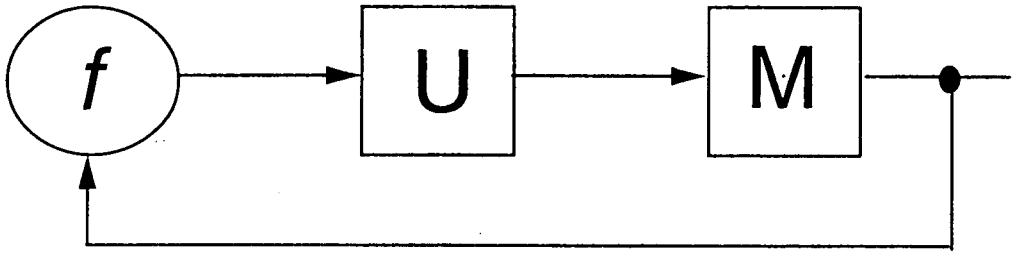

Figure 1

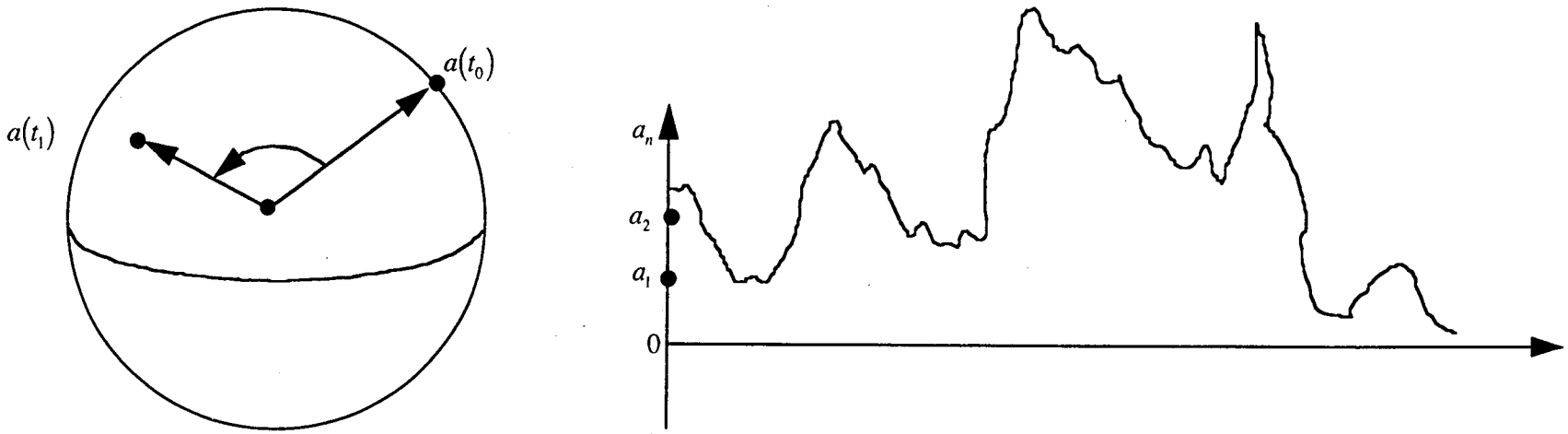

Figure 2a

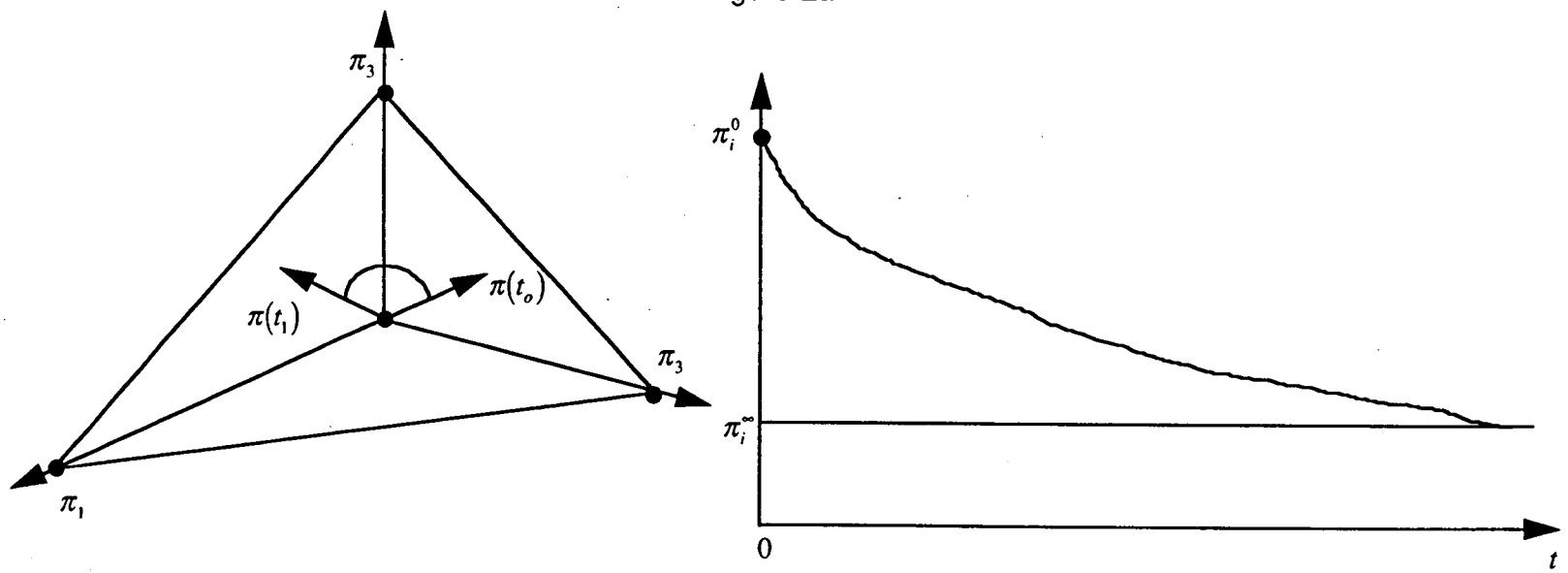

Figure 2b

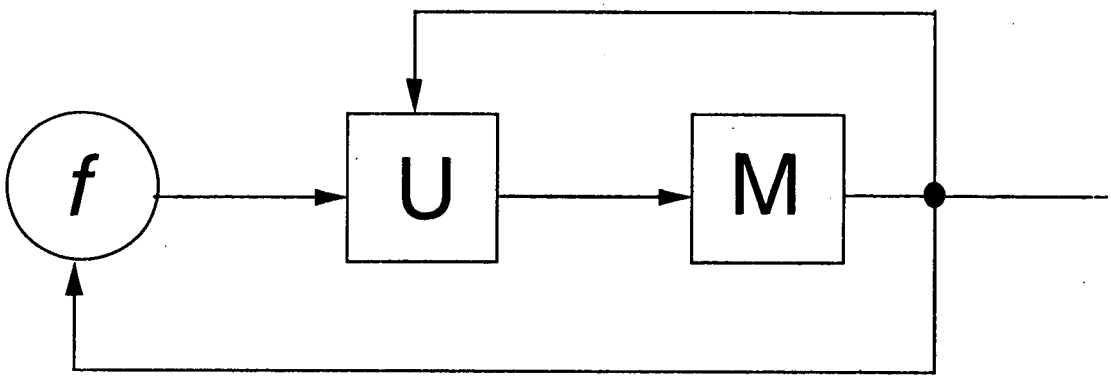

Figure 3

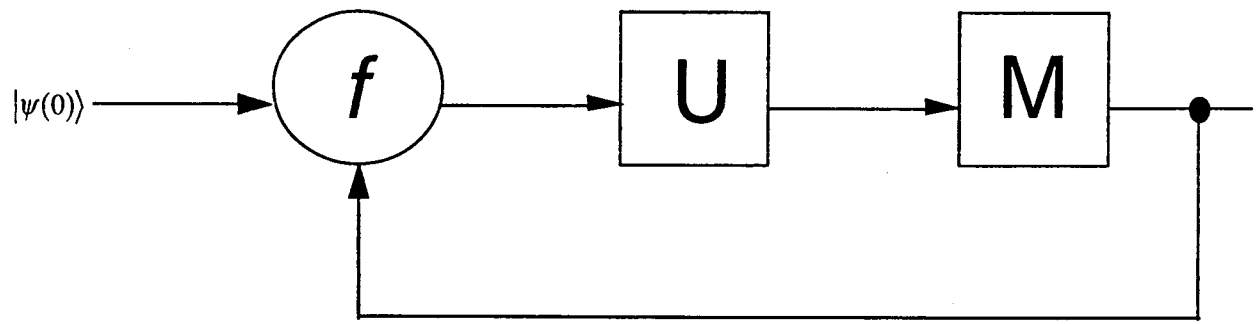

Figure 4

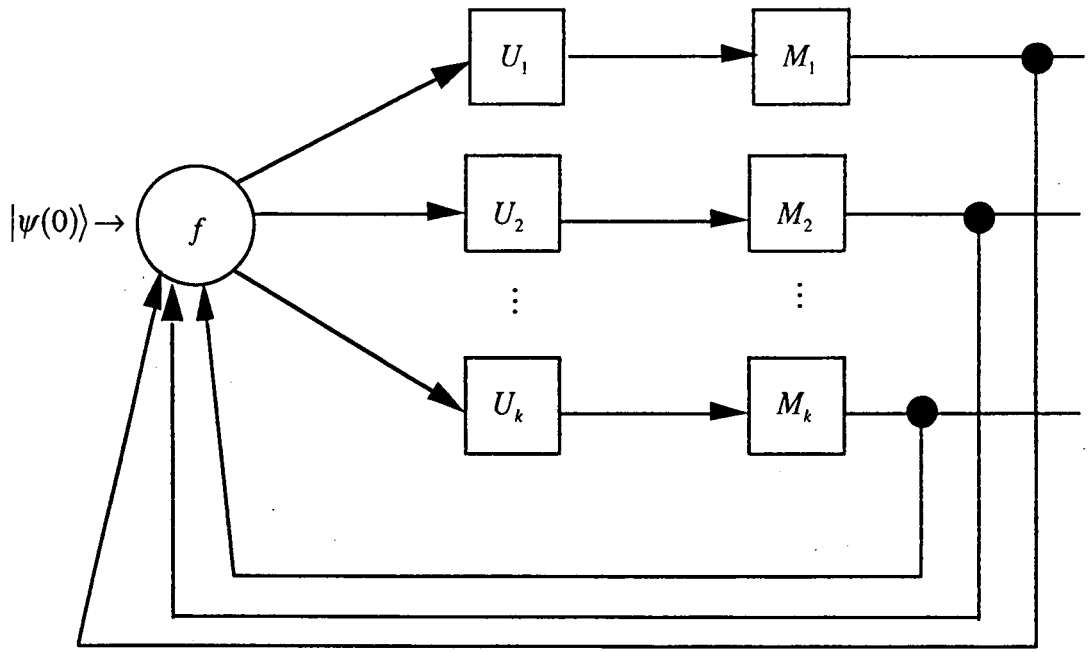

Figure 5

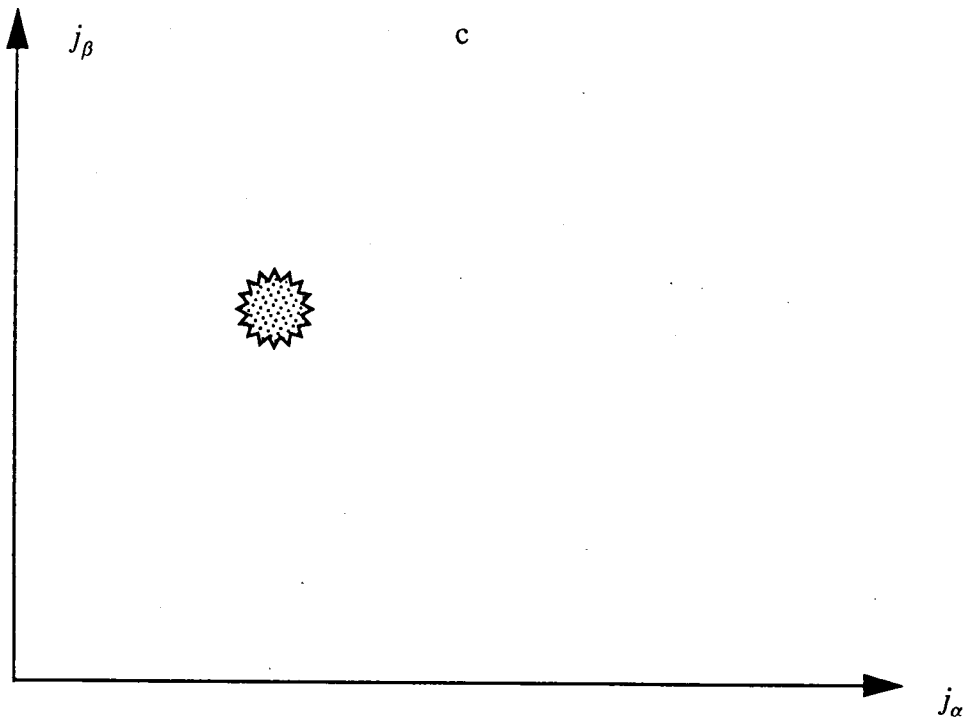

Figure 6

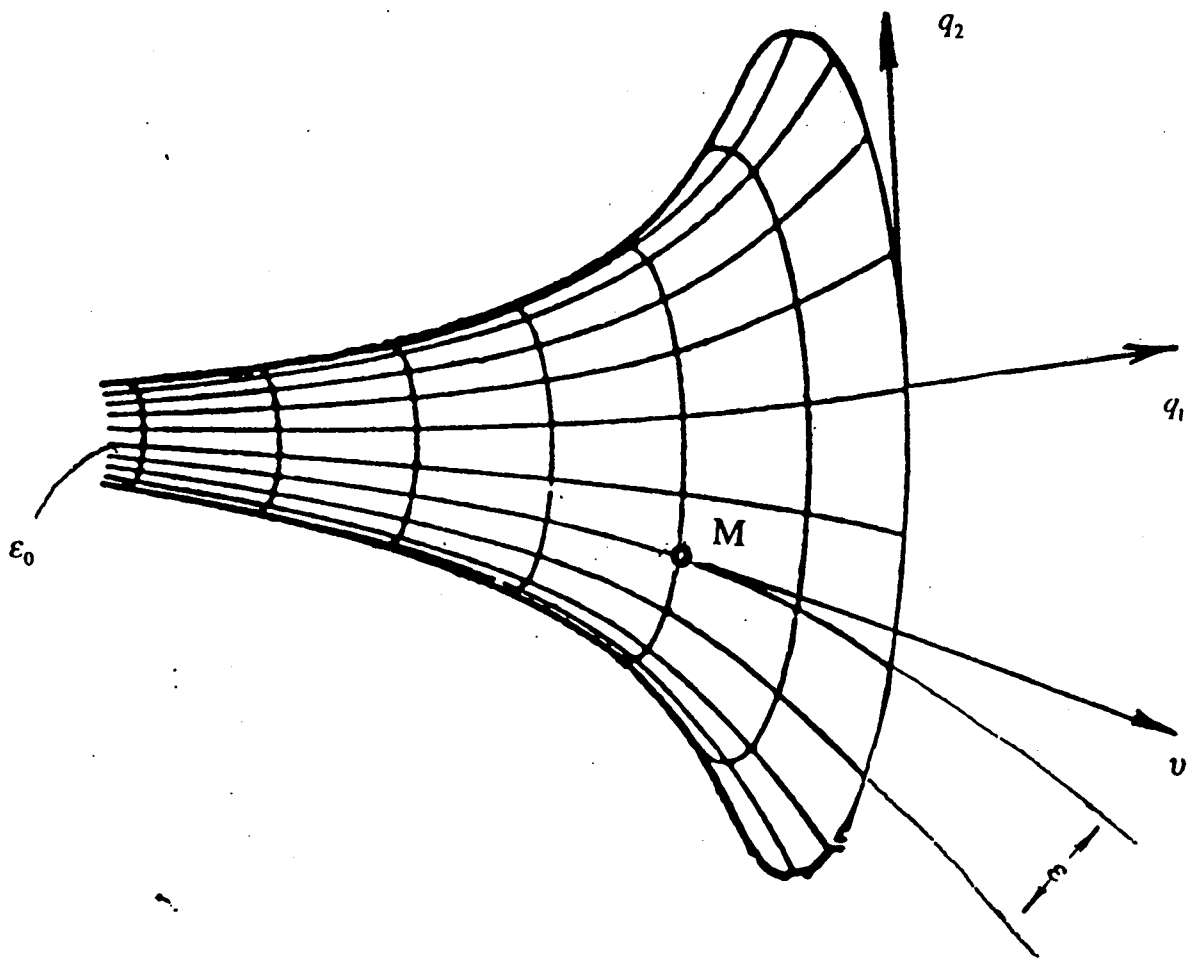

Figure 7